

Decision-theoretic relativity in deontic modality

Nate Charlow¹

© Springer Science+Business Media Dordrecht 2018

Abstract This paper explores the idea that a semantics for ‘ought’ should be neutral between different ways of deciding what an agent ought to do in a situation (e.g. different decision theories). While the idea is, I argue, well-motivated, taking it seriously leads to surprising, even paradoxical, problems for theorizing about the meaning of ‘ought’. This paper describes and defends one strategy—a form of Expressivism for the modal ‘ought’—for navigating these problems.

Keywords Deontic modals · Decision theory · Expressivism · Kratzer

Special thanks to David Plunkett. Thanks as well to Derek Baker, Fabrizio Cariani, Jennifer Carr, David Copp, Jamie Dreier, Stephen Finlay, Arno Goebel, Rory Harder, Dan Harris, Nathan Howard, Phil Kremer, Dan Lassiter, Sven Lauer, Robert May, Paul Portner, Jim Pryor, Adam Sennet, Eric Swanson, Will Starr, and audiences at the University of Konstanz, UC Davis, Dartmouth, Lingnan University, and New York University. The support of SSHRC (Insight Grant #435-2015-0423) is gratefully acknowledged.

✉ Nate Charlow
nate.charlow@gmail.com

¹ Department of Philosophy, University of Toronto, Toronto, Canada

1 Introduction

Theorists of the meaning of normative language—especially, though not only, deontic modals¹ like ‘ought’ and ‘should’—generally agree that a theory of its meaning should not bake in *substantive normative assumptions*. There is, in other words, general agreement that some version of a principle I will dub “Substantive Neutrality” should constrain semantic theorizing about ‘ought’.²

Substantive Neutrality (SN)

A theory of meaning for ‘ought’ should aspire to neutrality about substantive normative questions.

To give an initial sense of SN’s content, a semantics for the modal ‘ought’ should not, *inter alia*, entail that only ‘ought’-sentences consistent with consequentialism are true. Very roughly speaking, if (1) is true, it is not in any sense *analytic*.

(1) You ought to perform the action with the best outcome.

Not surprisingly, this “very rough” understanding has limited theoretical mileage: foundational matters concerning SN and its application to issues of great importance in theorizing about normative language are unclear. For instance:

- What is the basic motivation for SN?
- What makes a normative assumption “substantive” in the relevant sense?

¹ Linguists may think ‘deontic modal’ a poor name for this paper’s focus. 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) (cf. [Portner 2007](#)). Philosophers tend to use ‘deontic modal’ to cover roughly the same ground as the linguist’s ‘prioritizing modal.’ I follow the philosophers in this paper.

² The importance of a principle of normative neutrality for work on deontic modals owes largely to Jennifer Carr (see especially [Carr 2012, 2015](#)). Another key precedent is Gibbard (1990, Chapter 1). Recent literature on this question includes [Cariani \(2014, 2016\)](#) and [Lassiter \(2016\)](#). There is disagreement in the literature about whether Substantive Neutrality is an empirical or methodological constraint, about what makes a normative commitment “substantive” in the relevant sense, and so on. This paper will mostly prescind from these debates (though it will develop an understanding of what substantiveness, in the relevant sense, involves). The position developed here—in which judgments of semantic competence figure as the most prominent form of data—might be classified as either “methodological” (if one does not see such judgments as an empirical constraint on a semantic theory) or “empirical” (if one does). For what it is worth, I tend to regard such judgments as empirical in nature—indeed, obviously so—though I don’t believe anything should turn on this. In part, this is because judgments of semantic competence always seem to bring empirical commitments in tow. For example, if Ted knows what ‘I ought to go’ means, then a *de dicto* attitude ascription like ‘Ted thinks he ought to go’ should have a sensible truth condition. But, as Fabrizio Cariani has stressed, if a semantics for ‘ought’ makes a substantive normative assumption *N*, and Ted cannot be represented as accepting *N*, then it is hard to see how to state a sensible truth condition for the type of attitude ascription in question ([Cariani 2014, 2016](#)).

Similar principles to SN do appear in cognate literatures—one case is the philosophical literature on metaphysical modality.³ But the dialectic surrounding SN is special in certain striking ways (as I will argue). This essay will suggest that a core reason for valuing SN is a commonplace view of the general subject matter of semantic theorizing: *representing semantic competence*, i.e., what someone knows when they know the meaning of the linguistic item about whose semantics we are theorizing. This suggests a more determinate reading of SN. On one understanding of “substantiveness,” it is an intrinsic feature of a normative assumption’s *content*, so that normative assumptions are “substantive” when non-“trivial” in content. If, however, we read SN as being motivated by considerations of semantic competence, substantiveness will concern, not just the normative assumption’s content, but its role in paradigmatic (i.e. competence-dependent) instances of normative judgment and claim-making: a normative assumption is substantive, in the relevant sense, when it is false to attribute even a tacit commitment to it to each and every competent user of ‘ought’.

However anodyne in motivation, this understanding of SN generates striking, even paradoxical, effects downstream. Two of special interest: (1) certain “trivial”-sounding normative assumptions are plainly substantive in the intended sense; (2) the widely accepted theory of Kratzer (1977, 1981, 1991, 2012) achieves predictiveness by exploiting normative assumptions that count as substantive in the intended sense. So there is reason to think achieving even a modest degree of predictive power will mean violating Substantive Neutrality. We thus face an apparent dilemma: jettison the project of devising a predictive theory of meaning for ‘ought’, or jettison SN. Since the former horn must be avoided, it might seem as if we have the makings of an empirical case against SN and, more positively, for importing some of the fruits of, e.g., technical work in normative theory into the theory of meaning for ‘ought’—a strategy pursued notably by Lassiter (2011, 2016).

An alternative strategy—and the one ultimately pursued in this essay—is to outline a predictive theory of meaning for ‘ought’ that respects SN. Here we can make initial progress by treating normative language as semantically relative to a *normative perspective* (here understood as an agential representation of a salient decision problem and decision theory). What is the nature of this relativity? Is it analogous to ordinary forms of semantic relativity (e.g. context-sensitivity of a compositionally derived truth condition)? I argue that it is *not*: the semantic relation that relates a normative perspective to an ‘ought’ represents a cognitive relation—the relation of *competent acceptance*—rather than a relation that behaves anything like (relational) *truth* or *satis-*

³ Another, somewhat more ambiguous, case is meta-ethics. Many meta-ethicists would now agree that the theory of meaning for normative language should aspire not to bake in *substantive meta-ethical assumptions*. A semantics for ‘ought’ should not, for example, deliver moral realism as a corollary: if moral realism is true, its truth cannot be ascertained just by examining the semantics of ‘ought’. It is a virtue of the linguistically standard quantificational semantics for ‘ought’ (see especially Kratzer 1977, 1981, 1991, 2012) that it is compatible with a range of different meta-ethics (including Expressivist meta-ethics; see, e.g., Silk 2015; Charlow 2014b, 2015). Similarly, the sort of semantics for ‘ought’ articulated by, for example, Gibbard (1990, 2003) can be given a truth-conditional (e.g. contextualist) construal. This point has been noted often in recent work on Expressivism, which has tended to interpret it as a pragmatic or metasemantic, rather than semantic, thesis about normative language (Dreier 1999; Chrisman 2012; Yalcin 2012a; Silk 2013; Ridge 2014). Meta-ethical neutrality will be briefly addressed in Sect. 4.7.

faction in standard interpretations of model-theoretic semantics. The essay ends by outlining an *Expressivist* interpretation of this perspective-laden semantics for ‘ought’.

2 Clarification and motivation

SN is a familiar piece of advice for theorists of the meaning of ‘ought’. There are various reasons that might be offered in its support (and various understandings of SN accompanying these reasons). This section suggests a specific motivation for SN (and an accompanying, relatively specific reading of the principle).

2.1 Epistemic motivation for SN

A familiar—call it *epistemic*—motivation for principles like SN stems from general reflections on the appropriate methodology for any sort of theoretical inquiry.⁴ It is, in general, good theoretical practice to remain neutral about questions that do not bear directly on the thing one is theorizing *about*. Stalnaker (2011: 5–6) articulates this sort of position in writing about the methodology of theorizing about metaphysical modality:

[I]t is a virtue of a theoretical account of some concept [...] (a benefit in the cost-benefit analysis [i.e. the reflective equilibrium that characterizes theorizing about metaphysical modality]) if it is able to fashion some tools that manage to remain neutral on issues in dispute—to provide resources to formulate alternative

⁴ Familiar, that is to say, to philosophers of language and linguists. Meta-ethicists, however, will probably think immediately—I confess that I did not—of Gibbard (1990: Chapter 1). An evocative quote:

To refute an analysis [of the meaning of the word ‘rational’] by counterexample, we need a case in which a person not only doubts it, but accepts something inconsistent with it—without linguistic or logical confusion. Here is such a case: Octavia thinks reason demands that anyone give weight to his own future happiness. It makes this demand, she thinks, even on a person who is now indifferent to the future. Now whether or not she is right, if her thought is intelligible, if it is unconfused linguistically and logically, then the thesis is wrong as a claim about meaning. Future happiness is no goal of Cassius’s, imagine, and so by the Humean thesis [i.e. that one is rational just if one always does what one most prefers], it is rational for him to give it no weight. According to Octavia, he rationally must give it weight. If the Humean thesis is right as a claim about meaning, then Octavia has meanings wrong or else she is logically confused. If her opinion is intelligible, then the Humean thesis is wrong as a claim about meaning. (12)

Thanks here to David Plunkett. The view Gibbard sketches is similar in important respects to the view I defend here (and, like me, Gibbard’s central argument for a principle of Neutrality appeals to examples involving apparently semantically competent normative judgments that violate the Sure Thing Principle, on which more soon). Though we make use of similar examples to similar ends, there are important differences. Gibbard does not fully appreciate the *strong, positive empirical case* for the sorts of Neutrality-violations about which he is concerned, and as a result his analysis of normative language is non-responsive to it. (For the positive case, see Lassiter 2011, 2016 and this paper.) Further, the import of the conclusions Gibbard draws is limited to meta-ethical issues. This paper deals with the strong positive case for Neutrality-violations, and its major aim is to engage empirical questions that do not really appear on Gibbard’s radar. It also defends an Expressivist view of the role of normative decision theory in the semantics of ‘ought’ that avoids some of the difficulties that beset Gibbard’s semantics. For criticisms of Gibbard’s semantics, see Schroeder (2008a, b, c) and Charlow (2015).

substantive views as coherently as possible [...] I think we should be suspicious of an account of modality that tells us too much about what there is or about what there might have been.

Stalnaker means specifically to encourage methodological suspicion about theories of metaphysical modality that would commit us to things like talking donkeys or concrete counterfactual possibilities. Neutrality in Stalnaker's sense thus seems motivated by appeal to the *auxiliary epistemic burdens* incurred by competitor theories: theories (like Lewis' Modal Realism) that are not neutral tend to engender commitments—e.g., ontological commitments—that are controversial (in the sense that many theorists are liable to find them unacceptable). We are, moreover, liable to find these commitments unacceptable because they are unsupported by the kind of evidence we normally take to warrant such commitments; as Stalnaker writes, “I think most of us will find [these commitments] suspicious, not just because the benefit of the beautiful theory is outweighed by the cost of the ugly facts, but because there is something suspect about using this kind of theoretical virtue to reach this kind of conclusion about what there is” (6). To put it a bit blandly: it is bad methodology to take on controversial commitments in the course of philosophical theorizing, particularly absent appropriate independent evidence for those commitments.

2.2 Dominance

Intriguingly, an epistemic reading of SN seems to allow a theorist to make explanatory use of substantive normative assumptions in theorizing about ‘ought’. To see this, consider a case in which the relevant epistemic burden—whatever it looks like—is actually *met*. Suppose we have conclusive evidence—as we in fact do—that what one ought to do obeys a kind of *Statewise Dominance* constraint.⁵

Statewise Dominance (SD)

If, relative to a space of relevant contingencies $\{C_1, \dots, C_n\}$, α is preferred (conditional on C_i , for each i), then α is (unconditionally) preferred relative to $\{C_1, \dots, C_n\}$.

SD can be paraphrased using the modal ‘ought’ as follows⁶:

⁵ Statewise Dominance principles ordinarily receive formulations that are both stronger and more explicit than the versions invoked in this paper (cf. Joyce 1999: 17ff). I use relatively weak formulations to focus the discussion and our intuitions (since stronger versions introduce commitments that are orthogonal to the issues at stake here). I use relatively inexplicit formulations to show that Statewise Dominance principles must be further precisified in order to avoid counterexamples (as shall become clearer in the next section). For instance, a decision theorist might (and probably would) insist that the space of relevant contingencies $\{C_1, \dots, C_n\}$ invoked by Dominance principles form a *partition*. The apparent (and apparently paradoxical) incompatibility between the theorist's simultaneous *need for precisification* (to generate a normatively plausible Dominance principle) and need to *avoid precisification* (to avoid taking sides on substantive normative questions) is the central preoccupation of this paper.

⁶ On the tight link between statements of comparative preference and deontic modal claims, see Kratzer (1981) and discussion in Charlow (2016).

Modal Statewise Dominance (MSD)

If, relative to a space of relevant contingencies $\{C_1, \dots, C_n\}$, one ought (conditional on C_i , for each i) to α , then one ought unconditionally to α relative to $\{C_1, \dots, C_n\}$.

On a Stalnaker-inspired understanding of SN, a semantics for ‘ought’ is apparently permitted to make use of principles like MSD in accounting for apparent entailments between certain ‘ought’-sentences in natural language. For instance (supposing that the space of relevant contingencies is given by: {you wash your hands or wipe them on your pants, you wash your hands or use hand sanitizer}, and letting \checkmark indicate an intuitively valid inference):

- (2) a. If you wash your hands or wipe them on your pants, you ought to wash them.
- b. If you wash your hands or use hand sanitizer, you ought to wash them.
- c. \checkmark So, you ought to wash your hands.

Since the auxiliary commitment to (a principle like) MSD is independently well-supported, there is in principle no obstacle to accounting for (2)’s apparent validity in this way. What is more, the move *appears well-motivated*. It seems clear that a principle like MSD substantially underwrites the “good” feeling we get about the inference in (2): it will probably be the major part of an account of why this inference strikes us as impeccable.

Might we, instead, appeal to the meaning of the indicative conditional (and the “form” of sequence (2)) to underwrite the validity of this inference? It would appear not. Suppose we are thinking about the likeliest trajectory for a cyclone, and that the relevant contingencies are given by: {the cyclone moves North or South, the cyclone moves North or East}. We have estimates of the relevant probabilities at hand:

- $Pr(N) = .5$
- $Pr(S) = Pr(E) = .25$
- So: $Pr(N|N \vee S) = Pr(N|N \vee E) = 2/3$

Now consider (3), a minimal variation of (2), which is nevertheless degraded in comparison. (Note: ‘should’ here expresses a form of weak *epistemic* necessity concerning one’s expectations about the cyclone’s trajectory.)

- (3) a. If the cyclone moves North or South, it should move North.
- b. If the cyclone moves North or East, it should move North.
- c. #So, the cyclone should move North.

What plausibly *explains* the failure of (3) is the *falsity* of a corresponding *Epistemic Statewise Dominance* Principle (cf. [Yalcin 2010](#); [Lassiter 2011](#))⁷:

⁷ Depending on how one understands *validity*, the failure of (3) may suffice to render (2) invalid. To be explicit, my own view is that (2) is not valid, full-stop, but that it is semantically guaranteed in a pretty well-grounded sense (see the discussion of “*D*-validity,” which allows us to evaluate an inference for validity relative to a choice of decision theory, in Sect. 4.1). The problem with a semantics that violates Substantive Neutrality by assuming some precisification or other of MSD is that it will understand (2) as valid (full-stop) when it is not. I will explain why it is not appropriate to regard (2) as valid (full-stop) in a moment.

Epistemic Statewise Dominance (FALSE!)

If, relative to a space of relevant contingencies $\{C_1, \dots, C_n\}$, ϕ should (conditional on C_i , for each i) be the case (\approx for each i , ϕ is likely given C_i), ϕ should unconditionally be the case ($\approx\phi$ is unconditionally likely) relative to $\{C_1, \dots, C_n\}$.

It would seem obvious that (3) strikes us as degraded in comparison to (2) precisely because (3) tracks or instantiates a defective mode of reasoning by dominance.

2.3 Semantic motivation for SN

While Dominance reasoning substantially underwrites the “good” feeling we get about the inference in (2), it is, I’ll suggest, difficult to see how it could, in the end, appropriately underwrite an account of its *validity*. This shall motivate a different way of reading SN.

Note first that both SD and MSD were stated loosely and could (indeed, I will argue, must) be precisified to be put to theoretical use in, e.g., normative decision theory (or indeed in the compositional semantics of deontic modals). A helpful initial precisification restricts the application of dominance principles like these to spaces of contingencies that *partition* the relevant information. Notice, in particular, that Epistemic Statewise Dominance becomes *true* on such a precisification, and that the inference in (3) becomes impeccable if we imagine a context in which {the cyclone moves North or South, the cyclone moves North or East} is a partition of the relevant possibilities. Suppose Zeus has pressed exactly one of buttons A or B. If Zeus pressed A, the cyclone might move North or South (but likely North). If Zeus pressed B, the cyclone might move North or East (but likely North). The reader is invited to verify that the inference in (3) goes through, given such a context.

As this illustrates, the precisification of Dominance reasoning is—even in semantic theorizing—an essential prerequisite for putting it to plausible theoretical use. Once we start, however, it is hard to know where to stop. Here is a familiar illustration (cf. Joyce 1999: 115ff). Suppose you are deciding whether to give up smoking—an activity you greatly enjoy, but not enough to outweigh the possibility of a cancer diagnosis. The relevant contingencies are given by $\{\mathcal{C}, \neg\mathcal{C}\}$ (with \mathcal{C} the proposition that you get cancer) (Table 1).

Table 1 Should You Quit Smoking?

	\mathcal{C}	$\neg\mathcal{C}$
Quit	☹☹	☹
Don’t	☹	☹☹

In such a context, we would like to predict the truth of a sentence like (4):

(4) You ought to quit smoking.

However, since the relevant contingencies form a partition, MSD can apparently be invoked, to absurd effect:

- (5) a. If you're going to get [given that you get] cancer, you ought to smoke.
- b. If you're not going to get [given that you don't get] cancer, you ought to smoke.
- c. #So, you ought to smoke.

Some *further* precisification of MSD is evidently required to bar its application to this sort of case.⁸ (The usual solution is to restrict the application of Dominance reasoning to cases in which the relevant contingencies are *independent* of the relevant available actions.⁹)

There are, then, *parallel pressures to precisify* SD for decision-theoretic application, and MSD for semantic application. The central difficulty is that, when theorists attempt to precisify Dominance reasoning for such applications, they will arrive at different answers (plausibly because of previously unnoticed differences in certain “ground-level” normative commitments). The Sure Thing Principle of Savage (1972) is one famous example, as emphasized, to similar effect, by Gibbard (1990: 13ff) (see also note 9).

Sure Thing Principle (STP)

Given schematic decision problems P_1 and P_2 ¹⁰:

P_1	\mathcal{C}	$\neg\mathcal{C}$	P_2	\mathcal{C}	$\neg\mathcal{C}$
I	X	Z	III	X	W
II	Y	Z	IV	Y	W

S ought to select I, rather than II (\approx I is preferred to II) only if S ought to select III rather than IV (\approx III is preferred to IV).

STP is crucially understood by Savage as a *mathematical precisification* of the “mathematically useless” (i.e. imprecise or informal; 1972: 22) principle of suppositional practical reasoning that undergirds principles like SD.

What technical interpretation can be attached to the idea that [action] f would be preferred to [action] g if B were known to obtain?¹¹ [...] [T]he matter would

⁸ The dialectic here is worth clarifying. (5)’s (ersatz) legitimacy is overdetermined: it follows from MSD (as stated), as well as independently from the (alleged) validity of constructive dilemma. There is a very strong empirical case that constructive dilemma is invalid for natural language indicatives (see e.g., Willer 2012). Suppose we jettison it. The inference in (5) is still sanctioned by MSD (as presently formulated).

⁹ And of course the question of how to understand the independence of relevant contingencies on actions immediately raises the question of whether *independence*, in the relevant sense, is a *causal* or *evidential* notion (and hence immediately raises the normative question of whether rational agents should maximize Causal or Evidential Expected value) (see esp. Gibbard and Harper 1981).

¹⁰ In these tables, row names correspond to actions (alternatively, bets) available to the agent of the decision problem. Column names correspond to relevant contingencies (and typically partition some information space). Cells of the table correspond to outcomes associated with payoffs, utilities, or values; X, for example, is the payoff associated with choosing I given \mathcal{C} in P_1 (as well as with choosing III given \mathcal{C} in P_2).

¹¹ Savage mistakenly treats the (non-epistemic) ‘conditional on ϕ ’ construction invoked in SD and MSD as interchangeable with (epistemic) conditional constructions like ‘given that ϕ is known to obtain’. For why this is mistaken, see e.g., Weirich (1980). For why conditional preferences/utilities matter for theorizing about (conditional) deontic modality, see Charlow (2013a).

seem not to depend on the values \mathbf{f} and \mathbf{g} assume at states outside of B . There is, then, no loss of generality in supposing that \mathbf{f} and \mathbf{g} [...] [are] regarded as equivalent given $\neg B$ [...] If, after being modified so as to agree with another outside of B , \mathbf{f} is not preferred to \mathbf{g} ; then \mathbf{f} would not be preferred to \mathbf{g} , if B were known.

It is implicit in [this argument] that, if two acts \mathbf{f} and \mathbf{g} are so modified in $\neg B$ as to agree with each other, then the order of preference obtaining between the modified acts will not depend on which of the permitted modifications was actually carried out. Equivalently, if \mathbf{f} and \mathbf{g} are two acts that do agree with each other in $\neg B$, and $\mathbf{f} \geq \mathbf{g}$; then, if \mathbf{f} and \mathbf{g} are modified in $\neg B$ in any way such that the modified acts \mathbf{f}' and \mathbf{g}' continue to agree with each other in $\neg B$, it will also be so that $\mathbf{f}' \geq \mathbf{g}'$. (Savage 1972: 22)

But STP is famously called into question by the fact that the preferences of average persons seem to violate it in, e.g., the Ellsberg problem (Table 2).

[A]n urn contains 30 red balls and 60 balls that are either black or yellow in some unspecified proportion. In problem A, you are offered two options, I and II, and in problem B you are also offered two options, III and IV. A and B are distinct choice situations—you should not expect to play both games at once or in succession. In each case a ball is drawn at random from the urn and you receive specific rewards depending on the ball’s colour [\mathfrak{R} , \mathfrak{B} , \mathfrak{Y}] and your chosen option. The payoff table is as follows: (Steele 2007: 191)

Table 2 The Ellsberg Problem

A	$\mathfrak{R} [\frac{30}{90}]$	$\mathfrak{B} [\frac{k}{90}]$	$\mathfrak{Y} [\frac{60-k}{90}]$	B	$\mathfrak{R} [\frac{30}{90}]$	$\mathfrak{B} [\frac{k}{90}]$	$\mathfrak{Y} [\frac{60-k}{90}]$
I	100	0	0	III	100	0	100
II	0	100	0	IV	0	100	100

The “problem” arises from the fact that the (well-confirmed) majority preference in this sort of case apparently violates STP: most agents think they *ought to select I rather than II*, but also *ought to select IV rather than III*, even though this pair of preferences is plainly inconsistent with STP.¹² Some decision theorists have reacted to the fact that the majority preference in the Ellsberg problem is apparently STP-violating by rejecting STP as a precisification of the reasoning that underlies the sort of dominance reasoning Savage was interested in representing (e.g. Buchak 2013: Ch. 5). Others (indeed, Savage himself) retain STP, viewing the majority preference in the Ellsberg scenario as irrational.¹³

¹² Notice that I and III and II and IV are payoff-equivalent given $\mathfrak{R} \vee \mathfrak{B}$, and that I and II and III and IV are payoff-equivalent given $\neg(\mathfrak{R} \vee \mathfrak{B})$. A familiar, but informal, explanation for why most people have STP-violating preferences in the Ellsberg Problem invokes “risk” (or, more precisely, risk due to uncertainty): I is less “risky” than II (since one can be sure of the odds of \mathfrak{R}), but IV is less “risky” than III (since one can be sure of the odds of $\mathfrak{B} \vee \mathfrak{Y}$). While it is common to distinguish the type of risk exhibited by the Ellsberg Problem from the type of risk exhibited by the Allais Problem (see Sect. 3.4), I will not worry too much about that here.

¹³ Or else restrict STP’s application in cases in which it is impossible to assign a definite probability to the relevant contingencies. Joyce (1999: 101ff) and Steele (2007) provide a helpful overview of the dialectic.

Whoever is right, an extremely natural reading of these facts is that many people (not only theorists!) whom we take to be semantically competent with ‘ought’ *do not even tacitly accept* versions of Dominance that many people whom we also take to be semantically competent with ‘ought’ tacitly endorse. I will suggest that this is what makes any (sufficiently precise for theoretical applications) attempt to codify Dominance reasoning substantive, in the sense that is relevant for Substantive Neutrality. More generally:

Substantive Neutrality Precised (SNP)

A theory of meaning for ‘ought’ should aspire to neutrality about any normative question Q such that, for any answer A to Q , a semantically competent¹⁴ agent can fail to tacitly accept A .

If SNP is correct, the theory of meaning for ‘ought’ should avoid assuming any answer to this question which a semantically competent agent could fail to tacitly accept. *How to appropriately formulate the Dominance principle* that seems to underwrite the inference in (2) would seem to be precisely the sort of question that falls under SNP’s scope.

SNP is supported by a commonplace understanding of the subject matter of generative semantic theorizing. On this understanding, the semantics of, e.g., English is the study of what speakers of English tacitly know when they know the meaning of an English expression.¹⁵ If it is not reasonable to attribute even tacit acceptance of some normative view to each-and-every agent whom we count as semantically competent with respect to ‘ought’, that assumption counts, by SNP’s lights, as substantive, and it should not be baked into a theory of the meaning of ‘ought’. One small consequence of this—we will see others in due course—is that we cannot straightforwardly account for the apparent validity of (2) by assuming a semantics for ‘ought’ in which a specific Dominance principle (e.g. STP) is simply encoded.

¹⁴ I will not be providing an analysis of the notion of *semantic competence* in this paper. To forestall some questions: semantic competence comes in degrees: knowing that, e.g., *bicycle* refers to a physical object makes one minimally competent with the term; knowing that it refers to a physical object used in locomotion makes one relatively more competent; knowing that it refers to a physical object with two wheels makes one more competent still. Full semantic competence is *not* required for successful reference; relatively minimal competence is often sufficient. (A young child can use the term ‘bicycle’ to refer to a bicycle even if she knows only that bicycles are used by people to move around.) The amount of semantic competence required for full (or, better, sufficiently robust) competence is vague and, probably, context-dependent. The judgments of semantic competence and incompetence of which this paper makes use have, I think, the status of *data*. My reliance on them does not depend on understanding exactly what notion the term *semantic competence* ends up picking out.

¹⁵ Compare Yalcin (2014): “The notion of semantic value, as deployed in the context of [a general theory about what it is to speak and understand a language], serves fundamentally to characterize aspects of the mental state of knowing a language,” i.e., of what I have been calling “semantic competence”. This is a contested understanding of the subject matter of semantics, but it figures centrally in generative semantics (for further references and discussion, see Yalcin 2014). I will assume it here.

2.4 Neutrality

What kind of neutrality does Substantive Neutrality actually enjoin? A helpful test case is natural kind terms.¹⁶ Most theorists follow Kripke and Putnam in taking the chemical compound H₂O to figure centrally in the semantics of the natural kind term ‘water’; many, indeed, contend that the semantic value of this term is *exhausted* by its designation of H₂O. On the face of things, many speakers of English are semantically competent with this term, yet cannot be credited with even tacit knowledge of the nature of the chemical compound with which they brew their tea and quench their thirst. Thus there is a *prima facie* tension between Kripke–Putnam-style semantic externalism about natural kind terms and SNP—a tension which might seem to reflect poorly on the case for SNP.

We could, with some plausibility, dig in our heels and suggest that chemistry-ignorant users of ‘water’ are not in fact fully competent with the term. (Such speakers will, for example, tend to misidentify glasses *labeled XYZ* as water.) But here is a more conservative assessment: such speakers are competent in virtue of displaying competence with respect to a rule or concept that suffices—when combined with the relevant facts about their linguistic community—to fix the correct semantic value for ‘water’ in their mouths; they tacitly use this term to designate whatever substance plays the water-role in their linguistic community, and this is what their semantic competence with respect to ‘water’ ultimately consists in.¹⁷ There is ample linguistic knowledge to ground semantic competence for a chemistry-ignorant speaker with respect to ‘water’.

A principle of Substantive Neutrality about ‘water’—on which a theory of meaning for ‘water’ should aspire to neutrality about matters that are not relevant to semantic competence for ‘water’—is, on my understanding, compatible with a Kripke–Putnam-style semantic externalism about natural kind terms. Semantic competence with respect to ‘water’, I claim, consists in knowing a rule or possessing a concept that suffices—when combined with relevant facts about a linguistic community—to fix the correct semantic value for ‘water’. A speaker who, roughly, uses ‘water’ as we use it—and thereby displays the requisite semantic competence—tacitly accepts that ‘water’ is used to designate the substance her linguistic community uses it to designate: H₂O.

The question before us is this: could something similar be said for ‘ought’ and the various competing precisifications of Dominance reasoning? Let us try. The idea would be that some precisification or other of Dominance is *somehow settled* by the rule/concept, possession of which grounds a speaker’s semantic competence with ‘ought’, though possession of such a rule/concept does not, by itself, put someone in a position to know which precisification that rule/concept in fact settles.

But it is difficult to imagine a rule or concept that would both fix a sufficiently determinate formulation of Dominance while also being common to each speaker who is

¹⁶ Lassiter (2016) also appeals to an analogy with natural kind terms, but with the aim of undermining SNP. This section can be read as a response to Lassiter.

¹⁷ I am not endorsing any form of descriptivism about the semantics of ‘water’: I do not claim that ‘water’ has the meaning it has in virtue of having any kind of descriptive meaning, or in virtue of being associated by a speaker with a description that suffices, all by itself, to fix its reference. If anything I am assuming that speakers’ ability to refer with ‘water’ is parasitic on the ability of the linguistic community to refer.

competent with ‘ought’. I know of no evidence that the relationship between ‘ought’ and STP (e.g.) is analogous to the relationship between the relationship between ‘water’ and H_2O . Actually, there seems to be clear evidence that no such rule or concept is in place at all; surely this would be the most natural way of interpreting the facts put forward in the previous section.

An alternative explanation is, however, worth considering: perhaps present divergences over Dominance are analogous to divergences over the chemistry of water evinced by pre-Cavendish¹⁸ speakers of English: as they did not know the composition of the substance designated by ‘water’, perhaps we do not know which precisification of Dominance our use of ‘ought’ *in fact* settles. Consider an agent in the Ellsberg problem who has selected option I in decision problem **A**, and consider the following two pieces of conflicting advice for her in decision problem **B**:

(6) You ought to select III.

(7) You may select IV.

(6) is the sort of advice that would be offered by a defender of STP (e.g. Savage), while (7) is the sort of advice that would be offered by a defender of the majority “risk-averse” preference. The present suggestion is that exactly one of these pieces of advice is correct, though we *cannot know which* is correct until we settle the question of whether STP is a desirable precisification of the Dominance reasoning that seems to undergird (2).

This seems a rather fantastical suggestion—one which appears to be in tension with the available evidence. Consider the following case.

Cavendish

It is 1784; Henry Cavendish has just published the results of an experiment showing condensation of water when the contents of a vessel containing “common air” and “inflammable air” (i.e. hydrogen) are ignited. Johann and Antoine—both conversant in the relevant science, neither possessing any special expertise, Johann an advocate of the Phlogiston Theory, Antoine an opponent—are looking at a glass labeled XYZ. It is not yet fully clear what Cavendish’s experiments have shown: they are compatible with the idea that water is something *deposited* by “common air” when a mixture of common air and “inflammable air” is ignited, hence compatible with the notion that this glass of XYZ is water. Antoine follows Cavendish: the glass of XYZ cannot, he claims, be water because the underlying substance is not a *compound* of “inflammable air” and “common air”. Johann does not: the glass of XYZ is, he claims, water so long as it was deposited in the right way by ignition-stimulated phlogistication.

Johann and Antoine are each making assertions for which they lack sufficient evidence: neither yet knows quite what Cavendish’s experiments have shown (and therefore quite what substance ‘water’ really denotes). So it seems *both* are behaving

¹⁸ Henry Cavendish is the eighteenth century chemist usually credited with the discovery of water’s chemistry.

irresponsibly.¹⁹ For each speaker, it is quite possible that his assertion could turn out false: neither can say with well-founded confidence whether the glass of XYZ is in fact water, because neither can say with well-founded confidence what substance ‘water’ actually happens to denote. This is the sort of *semantic irresponsibility*—that is to say, epistemic irresponsibility grounded in a speaker’s failing to know (and knowing she fails to know) the semantic value of a term that figures centrally²⁰ in a claim she is asserting—we normally expect to detect in a case in which the semantic value of a term is not fully manifest to a competent speaker, but the speaker behaves as if it were.

Compare a context involving competing assertions of (6) and (7). Is there any semantic irresponsibility here to criticize?²¹ Speaking roughly, one type of agent uses (6)—entirely appropriately—to express one kind of normative commitment; the other uses (7)—also entirely appropriately—to express a distinct normative commitment. Each type of agent speaks correctly—in the minimal sense that each appropriately deploys semantic competence to express a view about what to do. The disagreement stems from different views about whether an STP-encoding way of decision-making *is correct*. There is no clear semantic irresponsibility in using claims like (6) and (7) to voice a (possibly tacit) commitment to a certain way of decision-making. Neither an agent who uses an ‘ought’ to express disagreement with the majority preferences in the Ellsberg problem, nor one who uses an ‘ought’ to express agreement with these preferences, is at risk of making the sort of mistake that Johann and Antoine make.

The defender of the ‘water’-‘ought’ analogy may reply that there is a disanalogy between these cases. Notice that, in the Cavendish case, a responsible speaker will *know that she fails to know* what substance the term ‘water’ actually happens to denote. In the case of (6) and (7), and supposing that the meaning of ‘ought’ does determine a way of decision-making, it is perhaps less plausible to think a responsible speaker will know that she fails to know what way of decision-making that is; perhaps responsible speakers are ignorant of the fact that the meaning of ‘ought’ does determine such a procedure! The suggestion, then, is that this sort of higher-order knowledge may

¹⁹ To be clear, I do not assume that the norm of assertion is knowledge (or any particular norm of assertion at all). I am noting that we can call the speakers’ assertions in this case irresponsible until the proper interpretation of Cavendish’s experiments is settled. There is no similar sense of irresponsibility in those cases that are often cited as counterexamples to the knowledge norm of assertion, e.g., the Prediction (Jack Aubrey) and Retrodiction (Sherlock Holmes) cases of Weiner (2005). Of course scientists often use assertions to responsibly advance *hypotheses* in cases where it is common ground that the hypothesis cannot yet be believed with any real confidence. Let us stipulate that this is not what is going on here.

²⁰ By a term *figuring centrally in a claim*, I mean something like: for the claim to have the truth-value the speaker takes it to have, the speaker must be right about what that term’s semantic value actually happens to be.

²¹ This argument does not strictly require the relatively strong claims in the main text here, i.e., that speakers of both (6) and (7) are behaving responsibly. It is enough for my purposes that at least one—perhaps the one that evinces a commitment to the correct decision theory—is behaving responsibly (whereas in the Cavendish case, neither seems to be behaving responsibly). Maitra and Weatherston (2010) observe an interesting related phenomenon: the correctness of a description of what agent *S* ought to do seems generally to be, all by itself, sufficient for “proper” assertion (by *S*) of that description of what *S* ought to do. I am not quite sure how to fit this observation into the story I am sketching here (in part because how the notion of *propriety* that figures in the literature on norms of assertion relates to the notion of *irresponsibility* at stake here is not entirely clear).

explain our sense of irresponsibility in Cavendish, while its absence may explain our lacking this sense in the case of (6) and (7), even though ‘water’ and ‘ought’ are semantically analogous in the relevant sense.

In reply: in the Cavendish case, if Johann *fails to know that he fails to know* what substance the term ‘water’ actually happens to denote, he is, I hope we are comfortable saying, making a clear mistake. Given the state of the science, he *should* know that he fails to know what substance the term ‘water’ actually denotes. A speaker of (6) or (7), however, need not be making a mistake of this sort. She may be making a *substantive mistake* (i.e., about what to do), if there is a fact of the matter about whether an STP-encoding way of decision-making is correct (see footnote 21). But she is at risk of making the same sort of mistake that Johann and Antoine make only if the meaning of ‘ought’ does somehow manage to determine a way of decision-making that settles the question of whether STP is correct. Our reticence in attributing such a mistake to (at least one of) the speakers of (6) and (7) seems to indicate that we see no real risk here, hence that we find it unlikely that the meaning of ‘ought’ would determine a way of decision-making that settles the question of whether STP is correct. Maybe we are wrong, but, once again, the evidential burden is on the theorist who sees things differently to say why.

Let us draw the following moral: a theory of meaning for term X of language L can in some cases (e.g. natural kind terms) make explanatory use of substantive truths about X 's subject-matter without risking impugning the competence of an ordinary speaker of L . In these cases, a speaker of L is typically committed—in virtue of the rules that govern the use of X in L —to these truths: we might say she accepts or is committed to them tacitly (or perhaps *de re*). In such cases, the theory of meaning is not “substantively neutral,” but that is because the rules governing the use of X in L are not ultimately substantively neutral. I have argued here that there is no evidence of rules governing the use of ‘ought’ in English that would suffice to settle which formulation of Dominance is correct. If there is a fact of the matter here at all—something I do not wish to take a stand on either way—it is nevertheless probably not settled by the rules governing the use of ‘ought’ in English.

2.5 Generalizing

I have used the Dominance principle as my main example, but what I have said will apply more generally. Dominance principles are not formulated in isolation: they are *consequences/predictions of normative decision theories*. A normative decision theory, for present purposes, is any object that instantiates the following structure.

Decision Theory (Schematic)

A decision theory \mathcal{D} is a pair $\langle \Pi_{\mathcal{D}}, \Gamma_{\mathcal{D}} \rangle$ such that:

- i. $\Pi_{\mathcal{D}}$ is a class of **decision problems** that are well-formed by \mathcal{D} 's lights.
- ii. $\Gamma_{\mathcal{D}}$ is a (possibly partial) **choice function** mapping a well-formed decision problem $\pi \in \Pi_{\mathcal{D}}$ into the set of things, $\Gamma_{\mathcal{D}}(\pi)$, that are permissible, by \mathcal{D} 's lights, for an agent in π to realize or bring about.

Roughly, a decision theory \mathcal{D} is a recipe for arriving at a range of \mathcal{D} -permissible alternatives in a well-formed (according to \mathcal{D}) decision problem an agent might face. Causal Decision Theory, for example, is such a recipe—one that exhorts the maximization of expected utility (as calculated using a certain kind of probability). But there are many more prosaic examples of decision theories, in the sense intended here: any recipe or instruction for transforming a *body of information* (roughly, the information relevant for the agent’s decision) and *body of preferences* (roughly, those priorities whose implementation is relevant for the agent’s decision) into a range of *permissible alternatives*, counts as a decision theory. In particular:

- Any agent whose decision-making system outputs an intention from the agent’s beliefs and desires can be understood to be implementing a decision theory.
- Judgments and utterances of (agent-directed) ‘ought’s tend to be generated by decision theories: one represents the situation of an agent, the information that is relevant in that situation, and the priorities whose implementation is relevant in that situation; and one (tacitly) employs a function—a decision theory—from such representations into a representation of the alternatives that would be permissible for the agent to realize in that situation.

There are, I insist, many different decision theories whose acceptance is consistent with full semantic competence with ‘ought’ (hence about which the semantic theory of ‘ought’ should remain neutral). I have already argued that certain decision theories that deliver STP as a consequence, as well as certain decision theories that do not, fall under this heading. Other decision theories that fall under this heading include: Causal Decision Theory (of the sort that says you ought to two-box in the Newcomb Problem); Evidential Decision Theory (of the sort that says you ought to one-box in the Newcomb Problem; here, see esp. [Gibbard and Harper 1981](#); [Joyce 1999](#)); risk-averse decision theories (e.g., ones that use MAXIMIN as their main decision rule; for a classic discussion of MAXIMIN within a specific form of political decision-making, see [Rawls 1971](#): §26ff); decision theories which attempt to take risk into account in other ways (e.g. [Buchak 2013](#)); knowledge-based decision theories and their cousins (e.g. [Hawthorne and Stanley 2008](#); [Neta 2009](#)); and many more besides these (including representations of the “decision theories” that ordinary agents use in generating verdicts about what to do from descriptions of the situation in which an agent is supposed to act).²²

3 Problems with neutrality

I have argued for a specific reading of the generally accepted constraint that a theory of the meaning of ‘ought’ should be neutral on substantive questions—namely, I claim, those questions whose various answers are consistent with semantic competence with ‘ought’. This section will outline some problems with this position. That it to say, it

²² Compare [Carr \(2012\)](#): “Our semantics for deontic modals shouldn’t preclude the possibility of expressing, without equivocation, the consequences of a reasonable decision rule” (171). I won’t rerun the argument for each of these theories—I hope the reader can see pretty easily that what I have said about different ways of encoding Dominance in a theory of rational decision-making should extend to each of these theories.

will try to build a kind of case against SNP (and thus to discharge a bit of the burden on the theorist who would encode a normative decision theory in the semantics for ‘ought’). The first problem is that the standard semantics for ‘ought’ in both linguistics and philosophy, Angelika Kratzer’s, is incompatible with neutrality: it encodes a normative decision theory. A much more serious problem is that, once we denude the semantics of such normative assumptions, we are left with a theory that is hard to recommend: the resultant theory seems to relinquish an unacceptable amount of predictive power.

3.1 The standard

Kratzer’s analysis of ‘ought’ (Kratzer 1977, 1981, 1991, 2012) is likely familiar to readers, so I will be brief. Kratzer treats specific readings of ‘ought’ as determined by contextually supplied resolutions of two distinct parameters: the **modal base** f and the **ordering source** g , each understood as supplying a set of propositions (when evaluated at some possible world).²³ The modal base supplies, roughly, a domain of relevant possibilities that are subsequently ranked by the ordering source. A deontic necessity modal quantifies (relative to a world of evaluation w) universally over the set of possibilities compatible with $f(w)$ that are best according to $g(w)$.²⁴

Of course, this semantics can only be predictive if there is some story about how to go from $f(w)$ and $g(w)$ to the relevant domain of quantification—else the domain of quantification is indeterminate and no concrete truth condition is generated. Kratzer supplies a story whose crucial bits tell the theorist how to transform an ordering source into a ranking on possible worlds (and subsequently a domain of quantification).

Definition 1 A possibility x is *at least as good as* possibility y by lights of $g(w)$, notation $y \leq_{g(w)} x$, iff $\{p \in g(w): y \in p\} \subseteq \{p \in g(w): x \in p\}$.

Definition 2 A possibility x is *strictly better than* possibility y by lights of $g(w)$, notation $y <_{g(w)} x$, iff $y \leq_{g(w)} x$ and $x \not\leq_{g(w)} y$.

Definition 3 Given modal base f , ordering source g , world of evaluation w : the *domain of quantification*, notation $\min_{\leq_{g(w)}}(f(w))$, is $\{x \in \bigcap f(w) \mid \neg \exists y \in \bigcap f(w): x < y\}$.

Unpacking: x is at least as good as y just if every priority satisfied by y is also satisfied by x ; x is strictly better than y just if every priority satisfied by y is also satisfied by x , but there is also some priority satisfied by x that y fails to satisfy as well. The domain of quantification is simply the set of possibilities compatible with the modal base that can’t be improved on (by lights of the considerations in g). We thus arrive at this truth condition for ‘ought’:

²³ Propositions in this literature tend to be understood as sets of possibilities (equivalently: functions from possibilities into truth values). This is a simplifying assumption on which nothing crucial turns.

²⁴ This makes some more simplifying assumptions—e.g. the Limit Assumption. Kratzer’s official truth condition for ‘ought’ is fairly complicated, but these complications turn out to be unnecessary so long as one assumes that the number of worlds ordered by \leq is finite. There are cases in which this assumption probably should not be made (see, e.g., Swanson 2008), but we will ignore them here.

Definition 4 *ought*(ϕ) is true at w given f and g , notation $\llbracket \text{ought}(\phi) \rrbracket^{f,g,w} = 1$, iff $\forall v \in \min_{\leq g(w)}(f(w)) : \llbracket \phi \rrbracket^{f,g,v} = 1$.

Putting the pieces together, the core idea is that it ought (given f and g) to be that p just if every possibility compatible with f that does a good enough job at achieving the priorities encoded in g is one in which p .

3.2 Against the standard

Kratzer's semantics has confronted a number of recent challenges, especially in the literature surrounding the Miner Paradox (Kolodny and MacFarlane 2010). Speaking generally and roughly, these challenges say that Kratzer's semantics *does not say enough* about how changes in information (of the sort induced when one entertains the antecedent of an indicative conditional) can affect the ranking that determines the domain of quantification for the deontic modal (Charlow 2013b; Cariani et al. 2013; Carr 2012, 2015; Silk 2014). This criticism can be glossed roughly as follows: Kratzer's semantics does not take the decision-theoretic phenomenon of *conditional desirability*—the fact that what is preferred can change given certain changes to one's information—properly into account (see footnote 11). If this is right, then, so far as decision theory is concerned, Kratzer's semantics encodes *too little*.²⁵

Here, I want to stress a rather different point: Kratzer's semantics does encode some decision theory or other (in the sense of Sect. 2.5), since it encodes a recipe for generating a range of "permissibles"—permissibly realizable alternative possibilities—from the decision problem characterized jointly by f and g . That decision theory is, moreover, substantive: there are decision theories seemingly consistent with semantic competence for 'ought', with which Kratzer's semantics is incompatible. So far as decision theory is concerned, Kratzer's semantics actually encodes *too much*.²⁶

²⁵ This debate has (not at all coincidentally) played out in a way that mirrors the structure of this paper. Kolodny and MacFarlane (2010) first noticed that the ordering relevant for interpreting deontic modals must be able to change given a contraction of the modal base, if we are to predict the consistency of:

- (8) a. The miners are in A or B.
 b. If they're in A, we ought to block A.
 c. If they're in B, we ought to block B.
 d. We shouldn't block A or B.

Charlow (2013b) and Cariani et al. (2013) criticized them for not saying *how* the ordering changed given such a contraction (and thus, inter alia, failing to predict that (8b) and (8c) were *true* in the Miner Puzzle), and offered alternative accounts of this change. Carr (2012, 2015) criticized these papers on grounds of neutrality: both papers invoked decision rules—Charlow (2013b) a knowledge-based decision theory, Cariani et al. (2013) MAXIMIN—to allow them to make these predictions. So there seemed to be a dilemma between predictive gaps and (decision-theoretic) neutrality violations. This is akin to the dilemma I am setting up here.

²⁶ Carr (2012, 2015) criticizes Kratzer (as well as Charlow 2013b; Cariani et al. 2013) for this reason too. However, according to Carr, Kratzer encodes the decision rule MAXIMAX, "the rule that one should choose the option that has some chance of having the best possible outcome. This is a straightforward consequence of the more basic commitment of that semantics: that we should always simply bring about the best possible outcome in the modal background" (177). But this question seems to me more complicated than this. As Kratzer (2012) notes, there is nothing impossible (or even odd) about an ordering source that demotes possibilities in which the *best* outcome (as determined by the portion of the ordering source that

Here, briefly, are some of the substantive decision-theoretic assumptions that Kratzer’s semantics seems to encode (compare Carr 2012, 2015; Cariani 2016).²⁷

- **Ordinal value.** It cannot represent cardinal differences in value: $g(w)$ induces an ordinal ranking on possibilities compatible with $f(w)$.
- **Qualitative information.** It distinguishes only between possibilities compatible with $f(w)$ and possibilities that are not. It cannot represent quantitative differences in probability or likelihood between such possibilities (and thus cannot represent the effect such differences tend to have on decision-making).
- **Incommensurability.** Whenever both $\{p \in g(w): y \in p\} \not\subseteq \{p \in g(w): x \in p\}$ and $\{p \in g(w): x \in p\} \not\subseteq \{p \in g(w): y \in p\}$, x and y are incomparable: $x \not\prec y$ and $y \not\prec x$. It thus fails to represent differences in priority between propositions in $g(w)$: it cannot straightforwardly represent the preferences of an agent S for whom x realizes one thing S cares about, and y realizes a distinct, incompatible thing S cares about, but x outranks y (e.g. because the priority x realizes is more important than the priority that y realizes).
- **STP.** Consider schematic decision problems \mathbf{P}_1 and \mathbf{P}_2 . Assume that the agent’s information and preferences over outcomes are fixed across \mathbf{P}_1 and \mathbf{P}_2 .

\mathbf{P}_1	\mathfrak{E}	$\neg\mathfrak{E}$	\mathbf{P}_2	\mathfrak{E}	$\neg\mathfrak{E}$
I	X	Z	III	X	W
II	Y	Z	IV	Y	W

Fix a world of evaluation w , and consider \mathbf{P}_1 . If from w all the best possibilities are possibilities in which I is selected—if, that is to say, \mathbf{P}_1 ’s agent ought to select I in w —then, since I and II are equivalent given $\neg\mathfrak{E}$, some X-possibility is better, by lights of $g(w)$, than any Y-possibility; very roughly, it satisfies strictly more propositions in $g(w)$ than any Y-possibility. Let x be any such possibility. Now consider \mathbf{P}_2 . In \mathbf{P}_2 , x also satisfies strictly more propositions in $g(w)$ than any Y-possibility. Since III and IV are equivalent given $\neg\mathfrak{E}$, some III-possibility (namely, x) satisfies strictly more propositions in $g(w)$ than any IV-possibility. Relative to a single $g(w)$, then, \mathbf{P}_1 ’s agent ought, in \mathbf{P}_1 , to select I only if that agent ought, in \mathbf{P}_2 , to select III.

These assumptions are encoded as matters of *definition*.²⁸ Ordinal value is encoded because the ordering source is taken to induce an ordinal ranking; qualitative informa-

Footnote 26 continued

tracks basic considerations of value) is realized by accident, chance, or luck, while promoting possibilities in which *second-best* outcomes are realized through skill, intelligence, or wisdom. (Whether this new semantics resolves the Miner Paradox is a different issue—one about which I am more pessimistic.) I do agree with Carr about Kratzer’s inability to represent cardinal value and probability, as we will see. I am here indebted to discussion with Fabrizio Cariani.

²⁷ It is not crucial that Kratzer’s semantics encodes all of these assumptions—only that it encodes some, and that the ones it does encode are objectionable in view of the proper understanding of Substantive Neutrality.

²⁸ The semantics thus seems to have the remarkable property that any object-language sentence expressing such a decision-theoretic assumption is in some sense a *theorem* of the semantics. It is worth noting that Kratzer’s semantics also encodes less normatively controversial assumptions like the Transitivity of \preceq , Acyclicity of \prec , and what is alternately known as Independence of Irrelevant Alternatives or Principle

tion is encoded because the modal “base” is defined as a qualitative body of information (i.e. a set of propositions); Incommensurability is encoded because \leq is defined in terms of \subseteq ; and STP is encoded because the domain of quantification $\min_{\leq_{g(w)}}(f(w))$ is defined independently of “global” characteristics of the decision problem. Given Sect. 2, however, it seems objectionable for a semantics for ‘ought’ to encode any of these assumptions. ‘Ought’-judgments that express the verdict of a cardinal utility function are not revelatory of semantic incompetence. The same goes for ‘ought’-judgments expressing verdicts of \mathbb{R} -valued probability measures; the utilization of priorities within one’s preferences that allow one to resolve prima facie incomparabilities; or STP-violating preferences. For all I have said, any of these things may make an agent practically or epistemically *irrational* or *mistaken*. But to implicitly conflate being mistaken with a form of semantic incompetence is itself a mistake.

3.3 Enriched ordering sources?

A natural maneuver for Kratzer’s semantics—and an independently common strategy for dealing with problems like those arising from the Miner Paradox—is to suggest that judgments that are in apparent tension with a semantic theory’s prima facie decision-theoretic commitments can be accommodated, by proposing nuanced resolutions of the modal base and (especially) ordering source (this is, broadly speaking, the strategy of [Dowell 2012](#); [von Fintel 2012](#); [Katz et al. 2012](#); [Charlow 2013b](#)).

To illustrate, consider again the Ellsberg problem. Note that it would be possible for Kratzer’s semantics to predict that (7) is true for an agent who has selected I, if III is (loosely speaking) “risky” and “risk” is regarded as a mark of disvalue in the context. In other words, the maneuver attempts to accommodate STP-violating preferences by, roughly, treating the proposition *that one minimizes risk* as part of the ordering source. To similar effect, we can construct ordering sources that resolve incomparabilities, by, for example, constructing “a derived ordering that encodes the notion that, other things being equal, *the more expectations that are satisfied, the better*,” so that the relevant ordering source induces a *total order*—ordering worlds according to the *number of propositions* in the non-derived ordering source that worlds satisfy—over worlds compatible with the modal base ([Katz et al. 2012](#): 493).

Here, this maneuver amounts to a change of subject. I am not suggesting that it fails for its intended target: accommodating canonical *object-language judgments* of truth, consistency, or entailment in, e.g., the Miner Paradox.²⁹ But the point I am making about neutrality is not a point about predicting phenomena in the object-language. It is a

Footnote 28 continued

α —i.e. if $x \in \min_{\leq_{g(w)}}(f(w))$, then x remains among the best worlds in any contraction of $f(w)$ containing x . Is implicit acceptance of such things a condition of semantic competence with ‘ought’? I am unsure: there are lively literatures in philosophy and economics challenging each of these assumptions; it is doubtful that the proponents of heterodoxy are, in virtue of their semantic competence with respect to ‘ought’, even tacitly committed to the assumptions in question. While I am inclined to attribute some irrationality to some participants in these literatures, attributing semantic incompetence with ‘ought’ could seem extreme. There may, however, be countervailing empirical considerations that will demand encoding such assumptions.

²⁹ I do, however, think the maneuver fails for its intended target (see [Charlow 2013b](#); [Carr 2015](#)).

point about the decision theoretic commitments of Kratzer’s semantics—commitments that we have accessed *directly*, i.e., by interpreting the semantics in our theoretical metalanguage, rather than *indirectly*, i.e., by interpreting the semantics through the lens of the predictions it makes about object-language phenomena. The theoretical interpretation rests only on the claims that (1) for deontic interpretations of ‘ought’, the relevant information (*f*) and relevant priorities (*g*) jointly characterize a decision problem, (2) a recipe for transforming a decision problem into a range of permissible alternatives counts as a decision theory. To interpret the claims of Sect. 3.2 as claims about the inability of Kratzer’s semantics to accommodate phenomena in the object-language is to misinterpret me.

To put the point somewhat differently, on the maneuver I’m here criticizing, it is, in the Ellsberg problem, possible that (7) is true for an agent who has selected I, if III is *risky* (and risk is regarded as a mark of disvalue). If, however, risk is regarded as disvaluable, the relevant decision table is in fact something like this (cf. Joyce 1999: 102) (Table 3):

Table 3 The Modified Ellsberg Problem

A	$\mathfrak{R} [\frac{30}{90}]$	$\mathfrak{B} [\frac{k}{90}]$	$\mathfrak{Y} [\frac{60-k}{90}]$	B	$\mathfrak{R} [\frac{30}{90}]$	$\mathfrak{B} [\frac{k}{90}]$	$\mathfrak{Y} [\frac{60-k}{90}]$
I	100	0	0	III	100^{+RISK}	0^{+RISK}	100^{+RISK}
II	0^{+RISK}	100^{+RISK}	0^{+RISK}	IV	0	100	100

But I stress: *I have made no claims about the predictions of Kratzer’s semantics for such decision problems.*³⁰ The claims I have made apply only to decision problems that are instances of the type invoked by STP. It is obvious that the Modified Ellsberg Problem is not of this type. Unless decision problems with the features I intend are, for some reason, excluded by considerations of semantic competence—but why would they be?³¹—it is hard to see how the observation that there are available alternative decision problems (that make palatable predictions about object-language phenomena) engages the concerns I am presenting here.

3.4 Denuding?

While any of the substantive assumptions of Sect. 3.2 could be replaced with something more substantively plausible, that would of course not really get to the heart of the problem: if SNP is taken seriously, a semantics for ‘ought’ should encode no

³⁰ Nor does STP make any predictions about whether preferring I to II commits one to preferring III to IV in such a decision problem. As Joyce (1999: 102) writes: “[T]he common ‘preferences’ in the Ellsberg problem can falsify [STP] only if they are *not* based on a desire to know the objective chances.” That is to say, if we understand risk-aversion in Ellsberg as a desire to know the objective chances in Ellsberg, the decision problem used to present the Ellsberg problem is misdescribed for risk-averse agents.

³¹ Perhaps semantically competent risk-averse agents confronted with the Ellsberg problem always end up psychologically representing a distinct decision problem. Similarly, perhaps, semantically competent agents who resolve prima facie incomparabilities between alternative possibilities do so by representing a distinct decision problem in which no incomparabilities are witnessed. It is important to recognize that these suggestions amount to quite strong (and very probably false) commitments in rational psychology; they certainly cannot be assumed a priori. For relevant discussion and references, see Kahneman (2003).

substantive normative assumptions, plausible or otherwise. Such assumptions should be excised.

What should go? Definition 1 to start: this definition is responsible for requiring the ordering source to determine an ordinal ranking (as well as for enforcing Incommensurability). Similarly, we should no longer assume that f and g are qualitative in nature: we should at least allow that the information relevant for evaluating an ‘ought’ can be probabilistic in nature, and that the priorities relevant for evaluating an ‘ought’ can characterize, e.g., a cardinal utility function. To avoid presupposing STP, we should no longer assume that the domain of quantification $\min_{\leq g(w)}(f(w))$ is defined independently of “global” characteristics of the decision problem; the recipe for generating $\min_{\leq g(w)}(f(w))$ should make room for decision-making that takes into account things like risk (without presupposing that risk-aversion must be represented by utilizing a decision problem in which risk is represented as a bad-making feature of an action).

What remains when these assumptions are excised? Something like the idea that ‘ought’ is a universal quantifier over a domain somehow determined—in a manner sensitive to global features of decision problems (like risk)—by a (possibly probabilistic) informational parameter f and prioritizing parameter g (that possibly induces a cardinal ranking of possibilities). We may still, however, gloss a sentence of the form *ought*(ϕ) as follows: if things go well enough, by the lights of the relevant information and priorities—where “well enough, by the lights of...” is understood so as to be compatible with a decision-rule that violates STP— ϕ must be true.

Alas we cannot be happy with the resulting decision theoretically-denuded account of ‘ought’. It cannot predict a great deal of what such an account should predict. The remainder of this section will identify some related aspects of this problem.³²

Indeterminate domains. The account on offer is unacceptably vague: f and g must determine a domain of quantification in a way that is compatible with semantically competent ‘ought’-judgments by agents whose decision-making violates STP. For the truth-condition of an ‘ought’ to be theoretically determinate, the domain must be generated as a *function* of f and g . No specification of the properties of this function seems to be forthcoming. Indeed, no specification of the properties of this function seems *possible*: how could a single function represent the decision-rule expressed by both STP-respecting and STP-violating ‘ought’-judgments?

Decision-theoretic content. The denuded semantics for ‘ought’ also cannot represent what is intuitively expressed by typical judgments and utterances of many ‘ought’-claims. When an agent judges an ‘ought’-sentence to be true (and subsequently asserts it), she typically³³ employs a decision theory (in the sense of Sect. 2.5). A decision theory is the thing that gets her from her representation of the relevant decision problem

³² The problem, to be clear, is not that this theory has *no* predictive content. It predicts, for instance, that ‘ought’ is a universal quantifier, and thus predicts some prima facie puzzling inferences involving ‘ought’ to be valid (for discussion, see von Fintel 2012; Cariani 2013). But it does not have enough.

³³ Many ‘ought’-judgments—perhaps “evaluative” *oughts* (cf. Schroeder 2011)—plausibly do not involve any sort of representation of a decision problem at all (and so we do not need to attribute use of a decision theory to the relevant agent in order to account for them). It may be that Kratzer’s theory is well-suited to representing the content of such judgments. But it will not do for the judgments I am discussing in the main text.

to a judgment of what ought to be done in that decision problem; as such, it is typically an essential “ingredient” in the content of a judgment to the effect that a specific ‘ought’-sentence is true. A semantics that represents only the *decision problem*—represents only an informational parameter f and prioritizing parameter g , but not the *rule by which f and g determine* a range of things the agent in the decision problem can permissibly realize—apparently fails to represent an essential “ingredient” in the content of the judgments in question.

Suppose we have two agents who agree about the characteristics of the relevant decision problem—whose judgments are apparently grounded in identical f and g —but who nevertheless disagree about what an agent in such a decision problem should do. Consider, e.g., the Ellsberg problem. Two speakers agree that in decision problem **A** an agent should select option I, but disagree about what to do in decision problem **B**. Recall examples (6) and (7), repeated here:

- (9) You ought to select III.
 (10) You may select IV.

The speaker who asserts (9) evinces her “use” of a decision theory that is consistent with STP, while the speaker who asserts (10) evinces her “use” of a decision theory that is not (e.g. a risk-sensitive decision theory). These speakers disagree, but not in virtue of disagreeing about any relevant feature of decision problem **B**.³⁴ We would like a semantics to *somehow* represent this disagreement. In denuding the semantics of decision-theoretic content, we risk preventing ourselves from doing this.³⁵

Truth and entailment. There is a strong suspicion that saying the right things about inferences (2) and (5) requires incorporating some kind of Dominance principle into our semantic theory. A semantic theory which declines to encode Dominance reasoning cannot straightforwardly account for either the apparent semantic guarantee associated with (2), or the special absurdity of (5) (which I am inclined to attribute to a particular misunderstanding of Dominance reasoning).

³⁴ One may suggest that this disagreement is located in the fact that such speakers evince different priorities—i.e. risk-indifferent and risk-averse—hence evince different g 's, hence do not actually agree on the relevant features of the decision problem. (This is akin to the earlier-mentioned suggestion—by those who have wished to defend the majority preferences in the Ellsberg problem while retaining STP—that the payoffs in the Ellsberg problem are misdescribed, since they fail to represent the “negative payoff” that is involved for the average agent in choosing a risky alternative.) I address this above. For further discussion, see Charlow (2016).

³⁵ This argument is not to be conflated with the Relativist’s well-known claim that Contextualism mishandles disagreement (e.g. MacFarlane 2011, 2014). My objection is *not* that the denuded Kratzer semantics cannot identify *propositional contents* expressed by (9) and (10) that are incompatible. I am not assuming—and would not wish to assume—that disagreement needs to be accounted for in such a way (for arguments against thinking of disagreement in this way, see Plunkett and Sundell 2013). I do assume that a semantics should represent such disagreement as disagreement *about something*: perhaps about a propositional content, but also perhaps about *which resolution* of a variable semantic parameter is appropriate to use in generating a judgment of the form *ought*(ϕ) (cf. Plunkett and Sundell 2013). A semantics that declines to represent decision theories has difficulty representing the disagreement associated with (9) and (10) as about anything at all.

Examples like this multiply. There is a strong suspicion that, if (9) is true, its truth is *due* to the correctness of a STP-respecting decision theory. Again, it is hard to see how to account for this relationship in a denuded Kratzer semantics.

To bring the point out, consider a case related to the Ellsberg problem—the Allais problem. A ball is drawn at random from an urn containing 100 balls (89 red, 1 black, 10 yellow), with the payoff to you as follows (Table 4):

Table 4 The Allais Problem

C	$\mathfrak{R} \left[\frac{89}{100} \right]$	$\mathfrak{B} \left[\frac{1}{100} \right]$	$\mathfrak{Y} \left[\frac{10}{100} \right]$	D	$\mathfrak{R} \left[\frac{89}{100} \right]$	$\mathfrak{B} \left[\frac{1}{100} \right]$	$\mathfrak{Y} \left[\frac{10}{100} \right]$
V	1,000,000	1,000,000	1,000,000	VII	0	1,000,000	1,000,000
VI	1,000,000	0	5,000,000	VIII	0	0	5,000,000

As with the Ellsberg problem, the paradox arises from the fact that the well-confirmed majority preference for this sort of case seems to violate STP: most agents think they ought to select V rather than VI, but VIII rather than VII.³⁶

On the face of things, there is something like an *entailment* between (11) and (12). (I will clarify the sense in which I take (11) to imply (12) below.)

- (11) If you select I in **A**, you ought to select III in **B**.
- (12) If you select V in **C**, you ought to select VII in **D**.

That is to say: if (11) is true, it is plausibly STP that makes it so. Further, STP’s responsibility for (11)’s truth should imply that there is a kind of semantic guarantee holding between (11) and (12): if one is true, that is due to STP, in which case the other must also be true. The denuded Kratzer semantics cannot account for this sort of apparent semantic guarantee. The denuded semantics represents only features of a decision problem: *f* and *g*; it does not represent the *rule by which f and g determine* a range of things the agent in the decision problem can permissibly realize. But decision problems **C/D** are utterly distinct from decision problems **A/B**: effectively the only thing they have in common is the fact that their payoffs are determined for each bet by drawing a ball from an urn. It is a mystery, on this account, why there would be any kind of semantic guarantee of the sort I have claimed exists between (11) and (12).

4 Competent normative judgment

It might seem that *any* theorist of the meaning of ‘ought’ is liable to confront the tension explored in the prior section. To give a theory that can cover every way ‘ought’ is used by a semantically competent speaker, the theorist needs to jettison any assumption not required by bare semantic competence with ‘ought’. But, to be predictive, a theory

³⁶ As in Ellsberg, a commonly cited explanation involves risk (though notice that the kind of risk-aversion on display in Allais is *not* the kind that is grounded in a desire to know the objective chances, as in Ellsberg). The risk in Allais is risk in a more ordinary sense of the word: preferring VI means trading a sure million for a 10% chance at 5 million (and a 1% chance of nothing). That’s risky! But preferring VIII to VII means trading an 11% chance at a million for a 10% chance at 5 million (which doesn’t strike us as all that risky).

needs to utilize a decision-theory. If we read this tension pessimistically enough, we may be inclined to conclude that the project of devising a predictive theory of the semantics of ‘ought’ is a hopeless one, and should be abandoned.³⁷ Alternatively, one could take the tension as a *reductio* of SNP, however plausible that principle might have seemed.

This section will sketch a more ecumenical alternative. It is possible to build a well-functioning theory for ‘ought’ around the following idea: the basic job of a theory for ‘ought’ is to represent the relation of *competent acceptance/judgment* between an agent and an ‘ought’. Such a theory deals smoothly with the tension I have just highlighted: it is predictive, while also remaining appropriately neutral on substantive (e.g. decision-theoretic) questions.

The proper interpretation of this semantics is, I will suggest, Expressivistic in nature: the fundamental relation of the semantics is one that holds between an *agent* and a *sentence*, in virtue of (non-propositionally-individuable) characteristics of that agent’s state of mind—a relation, that is to say, of *competent acceptance*. I will sketch a broad theory of normative communication in which this idea figures prominently. I will conclude by saying why the form of Expressivism in question should be regarded as neutral between various ways of resolving certain foundational issues in meta-ethics.

4.1 A general account

Let us start by considering a semantics for ‘ought’ in which utilitarianism (say) was hard-coded. There would be arguments against such a semantics that look quite like the arguments we gave against Kratzer’s semantics (Sect. 3.2), and hence some reason to revise the semantics, by excising this substantive normative assumption. How would we do that? The natural move would be to *introduce a parameter* into the semantics that somehow represented the role of considerations of value—thus allowing us to represent the ordering on possible alternatives that seems to be an essential part of a judgment of the form *ought*(ϕ)—without assuming that the desirability of an action was a function of how much happiness it brought about. (This is, roughly, the semantic role that *actually is* played by *g* in Kratzer’s semantics.) The general lesson is that, when the truth condition of ‘ought’ seems essentially to involve some object about whose properties we—as theorists—wish to remain neutral, we can achieve neutrality by introducing this object as a *separate semantic parameter*, while allowing the value of that parameter to vary as needed.³⁸

³⁷ One way to develop this pessimistic line of thought would be to assimilate the phenomena discussed in Sect. 3.4 to the inference in (2), while holding that *none* are properly semantic phenomena, since *none* can be given a properly semantic account without running afoul of Substantive Neutrality. My own view, as I will try to explain in the next section, is that all of these phenomena, including the inference in (2), can be given a semantic account (provided we take on the account I develop here). This is fortunate, since the phenomena in question do seem, on the face of things, to be semantic in nature.

³⁸ There is reason for caution—owing to concerns about overgeneration—about this mode of argument. There is, for example, controversy about the proper understanding what it takes for an action to be done ‘freely.’ Should we thus parametrize the semantic evaluation of the expression ‘free’ to different views about what it takes for an action to be done freely? That would be surprising if it were true. In reply, I claim

I submit, then, that a semantics for ‘ought’ should represent a *decision problem* π and *decision theory* \mathcal{D} —understood as a pair $\langle \Pi, \Gamma \rangle$ of a set of well-formed decision problems Π and a choice function Γ —as distinct parameters of evaluation.³⁹ Relative to some $\mathcal{D} = \langle \Pi, \Gamma \rangle$, the choice function Γ , when applied to some $\pi \in \Pi$, is the set of things⁴⁰ that are permissible, by \mathcal{D} ’s lights, for the agent of π to realize or bring about. Along these lines:

Definition 5 $\llbracket \text{ought}(\phi) \rrbracket^{\langle \Pi, \Gamma \rangle, \pi, w} = 1$ iff $\forall v \in \Gamma_w(\pi): \llbracket \phi \rrbracket^{\langle \Pi, \Gamma \rangle, \pi, v} = 1$.

Some clarificatory notes. This is presently intended as a semantics for the deontic ‘ought’ (though we will consider an extension to other modal flavors below). We allow that Π and Γ may vary in extension with a choice of w : which decision problems are well-formed at w can depend on which decision problems are available at w , and the particular properties of the choice function can, at least in principle, depend on w . We assume that $\llbracket \text{ought}(\phi) \rrbracket^{\langle \Pi, \Gamma \rangle, \pi, w}$ is *undefined* when $\pi \notin \Pi_w$.

This semantics makes no assumptions about what shape $\langle \Pi, \Gamma \rangle$ must take. In particular, Γ can be the choice function of *any* decision theory whatever—various Bayesian decision theories (like Causal and Evidential Decision Theory), qualitative (non-Bayesian) decision theories of the sort used by ordinary agents in decision-making, etc. In any case where $\Gamma_w(\pi)$ is defined, the semantics yields a concrete condition under which $\llbracket \text{ought}(\phi) \rrbracket^{\langle \Pi, \Gamma \rangle, \pi, w} = 1$.

The very best argument for introducing a semantic parameter in this way is that there are compositional mechanisms (e.g. binding or shifting by quantificational and intensional operators) that seem to require it. Though this is a complicated topic, there is evidence that compositional mechanisms require both decision-problem and decision-theoretic parameters. Since the former claim is less controversial—indeed, given my understanding of a decision problem, already implicit in Kratzer’s semantics⁴¹—I will provide some apparent examples of the latter:

- (13) If maximizing causal expected utility is the thing to do, you ought to two-box in the Newcomb Problem.

Footnote 38 continued

that we should do this if, but only if, there are empirical considerations that motivate this independently (akin to the empirical considerations discussed in Sect. 3.4, and the binding data discussed immediately below). I doubt that the empirical case can be established for ‘freely.’ I am optimistic that the empirical case can be made for, e.g., imperatives and other kinds of “normatively laden” language. I am not inclined to regard this as overgeneration; rather it is testament to the explanatory power of the theory given here. For related discussion, see Charlow (2016).

³⁹ This is also the strategy of Carr (2012, 2015). More on this below.

⁴⁰ What kind of things? Here I treat them as possibilities, but other options can be imagined. I am fond of the idea that $\Gamma_w(\pi)$ delivers a set of *actions* in contexts in which ‘ought’ has a practical (prescriptive, agent-directed), meaning. When $\Gamma_w(\pi)$ is a set of actions, a condition of the form ‘ $\forall v \in \Gamma_w(\pi): \llbracket \phi \rrbracket^{\langle \Pi, \Gamma \rangle, \pi, v} = 1$ ’ can be read as follows: for each action v permissible by \mathcal{D} ’s lights in π , ϕ is a way of doing v . I think this will be helpful in the project of distinguishing the so-called “practical” ‘ought’ from the so-called “evaluative” ‘ought’ (see Schroeder 2011). I will here abstract away from the complications this idea generates.

⁴¹ For examples of “binding” the informational dimensions of a decision problem, see Cariani (2016).

- (14) If you want to maximize causal expected utility, you ought to two-box in the Newcomb Problem.
- (15) To maximize causal expected utility, you ought to two-box in the Newcomb Problem.

Other theorists will surely try to explain these examples in some fashion other than what is being suggested here. Nevertheless, they are unsurprising—and admit of relatively straightforward explanation—if the function of the operators in question is as it appears to be: binding or shifting a decision-theoretic parameter of semantic evaluation.

4.2 Attractions

The theoretical attractions of this account are many. The clearest is that the semantics is compatible with effectively any normative commitment, decision-theoretic or otherwise. Nevertheless, it displays impressive predictive power. The apparently dominance-underwritten inference in (2) can be given a semantic account. We start by defining the following notion of validity—a notion more restricted than (full-stop logical) validity, but nevertheless obviously semantic in character.

\mathcal{D} -validity

Let $\mathcal{D} = \langle \Pi, \Gamma \rangle$ be a decision theory. The argument $\frac{\phi_1, \dots, \phi_n}{\psi}$ is \mathcal{D} -valid iff $\forall \pi, w: \llbracket \phi_1 \rrbracket^{\mathcal{D}, \pi, w} = \dots = \llbracket \phi_n \rrbracket^{\mathcal{D}, \pi, w} = 1 \Rightarrow \llbracket \psi \rrbracket^{\mathcal{D}, \pi, w} = 1$.

An inference is \mathcal{D} -valid, roughly, when the satisfaction of its premises (by \mathcal{D}) entails the satisfaction of its conclusion (by \mathcal{D}). So long as \mathcal{D} secures a respectable notion of Dominance, it is not hard to see the outlines of an account of (2)'s impeccability: for any such \mathcal{D} , the inference is intuitively \mathcal{D} -valid.

The inference in (5), by contrast, is easily rendered \mathcal{D} -invalid. Simply choose a \mathcal{D} on which Dominance reasoning is restricted to decision problems where the relevant contingencies are (causally or probabilistically) independent of the available actions.⁴²

By representing decision theories semantically, we avoid the various problems that afflict a decision-theoretically-denuded semantics (Sect. 3.4). We no longer fail to represent the essential decision-theoretic “ingredient” in the content of a judgment of the form *ought*(ϕ). We are thus able to represent the disagreement between the speakers of (9) and (10) as about something: the disagreement is about *which decision theory* to use in generating a judgment about what to do in decision problem **B**. (I'll say a bit more about this below.)

⁴² It's slightly more complicated than this: we need to pair an appropriate choice of \mathcal{D} with an appropriate semantics for the conditional (and appropriate statement of the consequence relation on which constructive dilemma is not generally valid). For discussion, see Willer (2012).

We straightforwardly account for the apparent semantic guarantee between (11) and (12). Fix an STP-respecting decision theory \mathcal{D} . It is not hard to see why $\llbracket(11)\rrbracket^{\mathcal{D},\pi,w} = 1$ would imply that $\llbracket(12)\rrbracket^{\mathcal{D},\pi,w} = 1$.⁴³ Again, the inference is \mathcal{D} -valid.

4.3 Modal flavor

Parametrization to decision theories is also the strategy of Carr (2012, 2015), though the formal details are different (and Carr is a Contextualist about the resulting semantics—a commitment I shall criticize below). The major difference is that Carr uses decision theories to induce an ordering on worlds in the manner of Definition 1: in Carr’s system, given a decision rule r , information state s , and value function v (assigning values in \mathbb{R} to worlds according to their degree of relative desirability), $r_{s,v}$ characterizes an ordering source (see Carr 2015: §5.4). For example, if r exhorts maximization of expected value, $r_{s,v}$ is the ordering source given by the singleton:

$$r_{s,v} = \{\text{you maximize } s\text{-expected } v\text{-value}\}$$

This $r_{s,v}$ induces an ordering on worlds in the standard way (Definition 1). Thus the domain of quantification for the deontic modal, relative to such an $r_{s,v}$, will tend to include only worlds in which you maximize s -expected v -value.

A prima facie reason for preferring Carr’s semantics to mine is that Carr’s semantics preserves the key insight of Kratzer’s original account, namely, that deontic modality is just a species of a systematically related genus of modalities (including modals that lack action-guiding or advisory readings, e.g., epistemic modals). On Carr’s account, like Kratzer’s, all modals take ordering source arguments (and characterize their domains of quantification in the manner of Sect. 3.1); deontic modals are distinguished by taking ordering source arguments that are generally resolved by appeal to a decision rule and $\langle s, v \rangle$ pair.

There is, I would like to suggest, no obvious barrier to using a semantics of the shape described in Definition 5 to represent non-deontic readings of the modal ‘ought’. I can only sketch my reasons for thinking such an extension can be effected in the barest outline here. I stress that this is not a worked-out account of how such an extension would go. It is only an attempt to develop the conceptual underpinnings of such an extension.

My semantics relativizes deontic modals to two non-standard parameters: a choice of decision theory and decision problem. Decision problems, in the decision theorist’s sense, are (roughly) tables comparing alternative actions in terms of their outcomes relative to different relevant contingencies. A *bare set of alternatives* (e.g. possible worlds that are, given what is known, candidates for actuality) might be understood as a decision problem, in an extended sense of the notion: not a set of alternative actions, but

⁴³ Note: the π superscript in $\llbracket(11)\rrbracket^{\mathcal{D},\pi,w}$ and $\llbracket(12)\rrbracket^{\mathcal{D},\pi,w}$ is assumed to be null, since the decision problems relevant for evaluating (11) and (12) are supplied explicitly by material in the antecedents and consequents of these conditionals. I again concede that, since these are conditionals (and conditionals are tricky), the details involved in actually securing this implication will not be trivial. But it is not hard to get a feel for how the general story would work: \mathcal{D} , if true, will subsequently guarantee the truth of both (11) and (12).

a set of alternatives for what the world might be like, given some relevant background information. In this light, a decision problem in the decision-theoretic sense might be regarded as a *more structured* analogue of a more general notion: one that presents not epistemic alternatives, but available alternative actions, relevant contingencies, and a way of modeling the outcome of performing an action in a relevant contingency. All modals that are evaluated against a background set of alternatives—so far as I can tell, this is all modals—might then be understood as relative to a (possibly structured) set of alternatives (hence as relative to a “decision problem”).

Similarly, all modals might be understood as relative to a decision theory, in an extended sense of that notion. A decision theory as I have presented it here is a structured object $\mathcal{D} = \langle \Pi, \Gamma \rangle$, with Π a set of decision problems well-formed according to \mathcal{D} and Γ a choice function determined by \mathcal{D} . There is, I would like to suggest, a sense in which any *selection function*—any device for selecting a set of alternatives from within a range of alternatives—might be regarded as (or, more precisely, might be regarded as encoding) a decision theory, in an extended sense, since its *functional role* in the semantics is the same as that of the choice function: to *winnow down the alternatives presented by the relevant decision problem*. From this perspective, a choice of *selection function* might amount to a choice of *decision theory*, in the extended sense of “decision theory” that I have suggested here.⁴⁴

Certain modal flavors (e.g., deontic) might be tied to certain ways of resolving the suggested relativity: deontic modals would be interpreted, on the view I am suggesting here, relative to a bona fide decision problem and decision theory (in the sense of Sect. 2.5). Other modal flavors (e.g., epistemic) would be interpreted relative to (1) a set of relevant alternative possibilities (what I here call a “decision problem” in an extended sense), (2) some recipe for selecting a subset of the set of relevant possibilities (what I here call a “decision theory” in an extended sense). Such an account, suitably developed, would represent one way of preserving Kratzer’s insight that the modal flavors should be modeled, roughly speaking, as species of the same genus. So far as this desideratum is concerned, I can see no reason, in principle, for preferring Carr’s semantics to mine.

Not, however, so far as Neutrality is concerned. It is important to note that Carr retains Kratzer’s problematic (Sect. 3.2) recipe for inducing an ordering over relevant alternative possibilities—I do not. In contexts where $r_{s,v}$ is, e.g., a two-membered set $\{p, q\}$, Carr predicts incomparabilities between worlds that satisfy p (but not q) and worlds that satisfy q (but not p)—incomparabilities that are problematic for reasons we’ve already seen. Such incomparabilities can, in principle, be resolved with added semantic structure (as seen in Sect. 3.3; see also Carr 2015: 706n34). But it is, I take it, preferable to simply avoid encoding the incomparability-generating decision rule

⁴⁴ Indeed, in Kratzer’s analysis, a modal base f is a set of propositions that, relative to w , characterizes a set $\bigcap f(w)$ of alternatives compatible with it, which are narrowed down to a domain of quantification by application of the selection function $\min_{\leq g(w)}(\cdot)$. I have here characterized the modal base f as a decision problem and the selection function $\min_{\leq g(w)}$ as a decision rule. Earlier, however, I characterized f and g jointly as a decision problem. Both characterizations are apt. Kratzer’s g is doing double-duty: g provides a list of criteria for preferring certain worlds to other worlds (not unlike a utility function) *as well as*, given the definition of $\min_{\leq g(w)}(\cdot)$ (Sect. 3.1), a procedure for selecting among alternatives compatible with $\bigcap f(w)$ (not unlike a decision theory).

that would necessitate this sort of repair strategy. Encoding such a decision rule is a violation of SNP; the prospect of added semantic structure to resolve incomparabilities does not change the fact that the semantics encodes a problematic decision rule (Sect. 3.3).⁴⁵

4.4 Representors

I say that \mathcal{D} and π **satisfy** *ought*(ϕ) (relative to w) when $\llbracket \text{ought}(\phi) \rrbracket^{\mathcal{D}, \pi, w} = 1$. What do I mean by this: what fact does the satisfaction of *ought*(ϕ) by \mathcal{D} and π represent? Ordinarily, the model-theoretic relation of satisfaction (e.g. of sentence ψ , by a possible world w) is understood as *truth* (of ψ , at w). Is this, in the end, a sensible way of thinking of the relationship between *ought*(ϕ), \mathcal{D} , and π ?

I am unsure. On the face of things, $\langle \mathcal{D}, \pi \rangle$ represents something quasi-psychological in nature: a normative *point-of-view* or *perspective* that an agent can instantiate. When $\langle \mathcal{D}, \pi \rangle$ satisfies *ought*(ϕ), this in turn seems to represent something like *acceptance*—by an agent representable as having normative perspective $\langle \mathcal{D}, \pi \rangle$ —of *ought*(ϕ). Maybe (as MacFarlane-type Relativists will perhaps think) it makes sense to say that a point-of-view—or a context of assessment associated with that point-of-view—can make *ought*(ϕ) true. Maybe (as I am inclined to think) it doesn't (cf. Charlow 2016).

We can, at this juncture, prescind from this question (though I'll say a bit more about it in Sect. 4.6). Provided that $\langle \mathcal{D}, \pi \rangle$ represents a normative perspective, proponents of each type of answer can, I suspect, agree on the following metasemantic position:

Representor Metasemantics

A semantic theory T for fragment L is a theory compositionally determining a satisfaction relation $\models_{L, T}$ between Representors R and sentences $\phi \in L$, such that: $R \models_{L, T} \phi$ iff, for each agent α such that R represents α , α competently accepts or believes that ϕ .

This way of thinking about the subject-matter of semantic theorizing seems to me the most natural way of understanding the semantics we began to develop in the last section. On this metasemantics, the job of a Representor is to represent a perspective. $\langle \mathcal{D}, \pi \rangle$ is one kind of Representor: an instance of the general type that is needed for understanding what it is to competently judge that a sentence of the form *ought*(ϕ) is true.

I have suggested that $\langle \mathcal{D}, \pi \rangle$ represents the normative perspective an agent has when that agent represents decision problem π and uses a decision theory representable with \mathcal{D} to decide what to do in π . A relevant metasemantic question is this: under *precisely what conditions* does the normative perspective $\langle \mathcal{D}, \pi \rangle$ represent agent α ? Generally, π represents an agent's decision situation, understood to include some representation

⁴⁵ Jettisoning the “ordering” component of Kratzer's “ordering semantics” might seem to put a theorist at a serious empirical disadvantage: orderings allow Kratzer (1981) to define a wealth of comparative relations between propositions and to exploit these in the analysis of gradable modal constructions in natural language. But Kratzer's understanding of these comparative relations is problematic to begin with (Yalcin 2010). Further, the introduction of decision-theoretic structure generally allows a theorist to define a strictly greater number of comparative relations between propositions (Lassiter 2011).

of, at least, relevant information (akin to Kratzer's f), relevant desires or goals (akin to Kratzer's g), and available actions. If an agent α represents her decision situation in the way encoded in π , then π represents α . \mathcal{D} represents an agent's decision-making procedure: the psychological mechanism responsible for transforming a representation of a decision situation π into a range of alternatives that the agent considers permissible in π . I assume that such psychological states are real, and that it makes sense to think of $\langle \mathcal{D}, \pi \rangle$ -type objects as representing them. These assumptions seem, *prima facie*, like reasonable ones.

4.5 Contextualism?

A Contextualist might offer an alternative spin on the semantics I have stated, on which the satisfaction relation is a (more or less) ordinary semantic relation. The idea for 'ought' would be something like this: '*ought*(ϕ)' is true at a context-world pair $\langle c, w \rangle$ iff the c -relevant $\mathcal{D} = \langle \Pi, \Gamma \rangle$ and π are such that every possibility compatible with $\Gamma_w(\pi)$ witnesses the truth of ϕ . \mathcal{D} and π can be treated as contextually-resolved arguments of the modal 'ought', yielding something like the following truth condition:

Definition 6 $[\textit{ought}_{\langle \Pi, \Gamma \rangle, \pi}(\phi)]^{c, w} = 1$ iff $\forall v \in [\Gamma]^{c, w}([\pi]^{c, w}): [\phi]^{c, v} = 1$.

The suggestion is most naturally read as a form of Indexical Contextualism: the context of utterance (perhaps together with supplementary information) fixes the semantic value of the allegedly context-sensitive expressions $\langle \Pi, \Gamma \rangle$ and π .

Whether this idea is workable will depend on the availability of a plausible corresponding *metasemantics*—an account of how the semantic values of the allegedly context-sensitive expressions $\langle \Pi, \Gamma \rangle$ and π are determined by context.⁴⁶ So let us ask: could context determine a decision theory and decision problem relevant for the evaluation of an 'ought'-sentence?

One suggestion is that the semantic values of the allegedly context-sensitive expressions $\langle \Pi, \Gamma \rangle$ and π are determined by the speaker. When, for instance, an STP-committed speaker asserts (9), what she says is glossed roughly as: *according to (my STP-respecting) decision theory, III is required*. But this is a surprising report of (9)'s content. This proposition is (or tends to be) obviously true: it is generally trivial that an STP-respecting decision theory requires that an agent who has selected I in decision problem **A** select III in decision problem **B**; this is exactly what STP says, after all. What would the point of asserting such a proposition as this be, in the typical context where it is obvious that the relevant decision theory is STP-respecting (and it is obvious what STP requires the agent of **B** to do)?⁴⁷

⁴⁶ This is a familiar point from the Contextualism-Relativism debate for epistemic modals. Recall the common charge that Contextualists have great difficulty saying *which body of knowledge* is relevant for the evaluation of an epistemic modal (in, e.g., Eavesdropping Cases). I will not engage this debate at any length here. I hope the things I say in the main text will nevertheless have the ring of plausibility.

⁴⁷ Plunkett and Sundell (2013) suggest a story on which the point of asserting such a proposition is to set the stage for negotiation about *what resolution* of the context-sensitive parameter to use in generating the relevant modal judgment. I actually agree that this is the right type of pragmatics for deontic modals. However, as we will see below, the supposition that the relevant 'ought'-sentence expresses a

A different metasemantics for the allegedly context-sensitive expressions (Π, Γ) and π could be pursued. Here, conveniently, are some suggestions from Carr (2015):

There are many ways that the decision rule parameter might be resolved. First, it can be given explicitly: for example, *If Maximax is right, then we should block one of the shafts*. Second, the contextually salient [decision rule] r in some contexts might be whatever the “one true decision rule” is. Conversational participants might tacitly presuppose an objectively correct decision rule, in the same way they might presuppose that there is an objectively correct body of moral norms—even when these rules or norms aren’t transparent to speakers. Finally, in a given context, which decision rule is relevant may be underdetermined. So which proposition is expressed by the sentences might also be underdetermined, and might get determinate truth values only supervalueationally. (707)

None of these suggestions, in my view, can form the basis for a workable metasemantics for the “decision rule parameter”:

- Decision rules are only seldom given explicitly.
- Contexts in which we find the participants *presupposing* that there is “one true decision rule” are rare (outside of the decision-theoretic literature). Why?
 - Such contexts would be expected to be defective, as a default. A speaker could not take her addressee to have sufficient information to *coordinate* with her on the referent of the relevant expression (Carr’s r , my \mathcal{D}).⁴⁸ This follows the more general fact that agents typically do not take themselves to even *have* a view about the referent of a phrase like “the one true decision rule.”
 - Such contexts would be pragmatically opaque. For reasons just seen, a speaker and her addressee will typically be unable to coordinate on *which proposition was expressed* by the speaker’s utterance. In the absence of coordination on the proposition expressed, I am at a loss for what could be communicated in such a context by a speaker’s utterance of an ‘ought’.⁴⁹

Footnote 47 continued

context-sensitive proposition is not required to access a pragmatics according to which utterances of deontic modals set the stage for negotiation and coordination around a normative perspective. Since the suggestion that the relevant ‘ought’-sentence expresses a context-sensitive proposition is otiose (indeed independently problematic), we might as well jettison it.

⁴⁸ For the view that context-sensitive expressions generally require speaker and addressee to *coordinate* on a semantic value for the expression, see King (2014) and MacFarlane (2016). As MacFarlane writes, “We have plenty of other flexible expressions, whose extensions are to a great extent up to the speaker to determine. The most obvious examples are bare demonstratives like ‘this’ and ‘that.’ In principle, I can use ‘that’ to refer to any object. But with this freedom comes great responsibility. I must provide my hearers with enough cues to enable them to associate my use of ‘this’ with the same object I do, or communication will fail” (7).

⁴⁹ Perhaps a *set* of propositions, each member corresponding to an admissible resolution of context-sensitive parameters (cf. von Fintel and Gillies 2011; King 2014)? (This may be one way of developing Carr’s supervalueational suggestion.) This suggestion may hold promise, if the speech act of asserting (or “putting into play”) a set of propositions can be somehow understood and/or integrated into extant pragmatic models. (Thankfully that task is not mine.) An attempt at providing a pragmatics for assertions of sets of epistemically modal propositions is given in von Fintel and Gillies (2011). I am doubtful that this pragmatics extends to the sorts of cases under consideration here.

All this is to say it is doubtful that a Contextualist interpretation of the semantics proposed above can be paired with an acceptable metasemantics. The next section sketches an account of communication that avoids issues like these altogether.

4.6 Communication

I have claimed that ‘ought’-sentences are semantically relative to normative perspectives, and that this relativity is not analogous to the sort of relativity observed in context-sensitive expressions (i.e. the contribution of a propositional constituent as a function of the context of utterance). What is the nature of the relativity I have in mind? How does it connect to an account of what agents are doing when they assert ‘ought’-sentences?

Here is one answer (which I do not favor⁵⁰) (cf. Yalcin 2012b):⁵¹ the satisfaction relation between a normative perspective and an ‘ought’-sentence is indeed a truth-like relation. The point of asserting an ‘ought’-sentence is to express its *content* (with ‘content’ understood in the usual way, as the set of indices at which the sentence is true); asserting an ‘ought’-sentence expresses, roughly, the *set of normative perspectives* relative to which that sentence is true. (Note that, for simplicity, we are suppressing the semantic role of the world-index in this section.) The point of expressing such a content is to provide information in a familiar sense, roughly by narrowing down the list of candidates for “actuality”: given an input set \mathcal{N} of normative perspectives, typically the intersection of \mathcal{N} with the content of an ‘ought’-sentence will be a strict subset of \mathcal{N} . (More or less) psychological properties like *accepting STP* are subsequently given a supervaluational interpretation: a set \mathcal{N} of normative perspectives accepts STP iff STP is a property of (the decision-theoretic member of) each normative perspective in \mathcal{N} .

Here is a more neutral answer: an agent who asserts a sentence of the form *ought*(ϕ) is simply *expressing (competent) acceptance* of *ought*(ϕ). More carefully, such an agent expresses a competent judgment holding, in general, of an agent α iff α has a Representor $\langle \mathcal{D}, \pi \rangle$ such that $\langle \mathcal{D}, \pi \rangle$ satisfies *ought*(ϕ). In other words, such an agent expresses a content representable as a *property of Representors* (and, derivatively, a property of the agents whom those Representors represent), namely $\llbracket \text{ought}(\phi) \rrbracket$:

$$\lambda \langle \mathcal{D}, \pi \rangle \cdot \llbracket \text{ought}(\phi) \rrbracket^{\mathcal{D}, \pi} = 1$$

Such a property of Representors is, of course, equivalent to a set of Representors. But, as we will see, thinking of the semantic values of ‘ought’-sentences as *properties*, rather than as sets of candidates for “actuality”, allows us to state a pragmatics that prescind from certain rather difficult questions (e.g., whether or not it makes sense to think of a set of normative perspectives as a set of candidates for “actuality”).

⁵⁰ Charlow (2015) argues that strategies like this have difficulty with a version of the Frege–Geach Problem.

⁵¹ Willer (2013) gives a story like Yalcin’s, but states it in a dynamic framework. The presentation here can be easily restated in dynamic terms.

To see this, let us attempt to address the question of what an agent is doing when she asserts an ‘ought’. Why do agents express judgments whose content is a property of Representors? Consider a context in which two agents disagree about what a third agent should do in decision problem **B** of the Ellsberg problem. (Suppose the third agent has already selected I in problem **A**.) One asserts (16):

(16) You ought to select III.

This agent expresses, I have suggested, her acceptance of a normative perspective individuable with the property $\llbracket(16)\rrbracket$. Roughly, then, the agent expresses her acceptance of a decision theory that respects STP, knowing that her interlocutor, who accepts (17), rejects such a decision theory.

(17) You may select IV.

On my view, these utterances make it manifest that the interlocutors disagree on the relevant psychological property: there is no \mathcal{D} that satisfies the properties $\llbracket(16)\rrbracket$ and $\llbracket(17)\rrbracket$ (indeed, given Definition 5, these properties are logically incompatible). It thus sets the stage for *negotiation* and, ultimately, *coordination* about the relevant psychological property.⁵² There is much more to say here (cf. Charlow 2015), but this is a fine start.

This general account is compatible with the Yalcin-style account presented above. On a Yalcin-style account, $\llbracket(16)\rrbracket$ is a set of normative perspectives. The point of expressing $\llbracket(16)\rrbracket$ is to propose that a relevant set of normative perspectives \mathcal{N} be intersected with $\llbracket(16)\rrbracket$. If the proposal is accepted, \mathcal{N} is updated to \mathcal{N}' , with $\mathcal{N}' = \mathcal{N} \cap \llbracket(16)\rrbracket$. In which case, $\mathcal{N}' \subseteq \llbracket(16)\rrbracket$, and, given the intended supervalueational interpretation of psychological properties, *STP is accepted* relative to the updated state \mathcal{N}' . If the proposal cannot be updated on—as must be the case, if \mathcal{N} is the set of normative perspectives compatible with the interlocutor’s beliefs, given that the interlocutor accepts (17)—the fact that the proposal was made sets the stage for negotiation and, ultimately, coordination about the psychological property of accepting STP. In either event, the intended “upshot” of the utterance is coordination around the psychological property of accepting STP.

The general account is, however, compatible with a less committal pragmatics. On the story I favor, $\llbracket(16)\rrbracket$ is a (more or less) psychological property. To utter (16) is to express that property, with an eye to negotiation and, ultimately, coordination around it. The story *could* be supplemented with a “Dynamic Pragmatics” (Portner forthcoming) modeling how, precisely, coordination is achieved in certain kinds of cases (namely, cases in which the normative perspective targeted for update is compatible with the property in question); indeed, this is just what the Yalcin-style account provides. My own view is that formal models saying how, precisely, “new” information is integrated into “old” information are part of *diachronic normative epistemology*, and

⁵² This language is similar to that invoked by Plunkett and Sundell (2013). As I have indicated, I am sympathetic with a good deal of what they have to say about the pragmatics of disagreement.

cannot be part of linguistic theory, on grounds of Neutrality (cf. Charlow 2014a). In a slogan: semantic and pragmatic theorizing furnishes a theory of cognitive directives. Normative epistemology furnishes a substantive theory of diachronic compliance for cognitive directives. This is reason to avoid supplementing the general pragmatic account I have given with a precise account of how coordination is achieved in the relevant class of cases.

Whoever is right—a matter I cannot settle here—there will be a workable story of what agents are doing when they assert ‘ought’-sentences waiting in the wings.

4.7 Expressivism

The general story I developed here draws heavily on what Expressivists have tended to say about these matters (see Gibbard 1990, 2003 and discussion in Yalcin 2012a; Charlow 2015). The clearest hallmark of the account’s Expressivism is its demotion of propositional content in semantic and pragmatic theorizing for deontic modals: (non-propositionally individuable) psychological properties are doing all of the interesting semantic and pragmatic work. It is thus appropriate to label the interpretation I have offered of the semantics of Sect. 4.1 an Expressivist interpretation. It is, no doubt, surprising that this sort of Expressivism would emerge from the empirical and methodological reflections on the semantics for ‘ought’ that have occupied us here. Surprising or not, this does seem to me to be the best way to understand how things have gone here.

A final question I wish to address is whether the Expressivist proposal developed here is in fact neutral on relevant normative questions. A common (and commonly rebutted) charge is that Expressivism is a form of *nihilism* about the normative: if normative discourse is non-propositional, then we cannot make sense of the notion that normative claims are objectively true or false (since ordinarily a claim is true at a context of utterance c just when it expresses a proposition that is true relative to $\langle c, w_c \rangle$). But surely the question of whether or not normative claims are objectively true or false is a substantive normative question!

In reply, it suffices to note the availability of different *metaphysical postures* toward the semantics on offer here. The metaphysical posture appropriate to the commitment that normative claims are objectively true or false is, roughly, this: certain Representors count as *metaphysically privileged* when it comes to evaluating ‘ought’-claims, in virtue of representing a point-of-view that tracks *the way things really, actually, verily are (normatively)*.⁵³ An ‘ought’-sentence is *really, actually, verily true* just when it

⁵³ How can some $\langle \mathcal{D}, \pi \rangle$ track the way things really, actually, verily are, normatively? I envision an advocate of this view saying something like the following: an agent who is representable with $\langle \mathcal{D}, \pi \rangle$, inter alia, represents \mathcal{D} as the correct decision theory for deciding what to do, given π . Such an agent tracks the way things really, actually, verily are, normatively, only if \mathcal{D} really, actually, verily is the correct decision theory for deciding what to do. Expressivists of a Deflationary stripe frustratingly insist on asserting things like *there is a way that things really, actually, verily are normatively*. This complicates the project of distinguishing different meta-ethics by distinguishing between different metaphysical postures toward Representors. I do assume that there is *some* metaphysical criterion that will allow us to distinguish an Expressivist/anti-Realist posture toward Representors from a Realist posture toward them, even if I have not hit on exactly the right criterion here.

is evaluated as satisfied relative to a Representor that tracks the way things really, actually, verily are (normatively).

We are, of course, generally willing to think that there is a way that things actually are—that there is a privileged choice of parameters of semantic evaluation—when it comes to prosaic matters of fact. A sentence like (18) is *actually false*, i.e., relative to any choice of indices we take to represent actuality, though it *could have been true*, i.e., relative to a choice of indices we do not take to represent actuality.

(18) Faulkner never completed *As I Lay Dying*.

Nevertheless, a speaker can competently believe and assert (18), in virtue of being related by belief and assertion to a semantic content—standardly, the proposition that Faulkner never completed *As I Lay Dying*—that, while actually false, would have been true, relative to a non-actual choice of index.

My suggestion here is simply to transpose these bits of received wisdom to the non-standard parameters of semantic evaluation for which this paper has advocated. Example (18) illustrates that there is nothing fishy *per se* about the notion of a metaphysically privileged Representor (qua index or parameter of semantic evaluation). As ever, the notion of *actual truth* may be defined as truth (or satisfaction) relative to some class of indices which we have independent reason to think about as metaphysically privileged. A speaker can nevertheless competently believe and assert a claim that expresses commitment to a false decision theory—e.g., that one ought to One-Box in the Newcomb Problem—in virtue of being related by belief and assertion to a semantic content—roughly, on my account, the property a decision theory has when it recommends One-Boxing in the Newcomb Problem—that, perhaps, is actually false of a privileged class of (Causal) Decision Theories, but could have been true, i.e., relative to a choice of (Evidential) decision theory which we do not take to represent normative actuality.

Whether one thinks there is a way that things *actually are normatively*—hence whether one thinks certain Representors are metaphysically privileged when it comes to evaluating ‘ought’s—will, however, depend on one’s extra-semantic meta-ethical commitments. Certain kinds of Moral Realist are likely to be comfortable with this commitment; certain kinds of Moral Antirealist—this author, for instance—are not. Meta-ethicists of many different stripes may, then, enjoy the empirical and theoretical benefits of the Expressivism about ‘ought’ I have defended here.

References

- Buchak, L. (2013). *Risk and rationality*. New York: Oxford University Press.
- Cariani, F. (2013). ‘Ought’ and resolution semantics. *Noûs*, 47, 534–558. <https://doi.org/10.1111/j.1468-0068.2011.00839.x>.
- Cariani, F. (2014). Attitudes, deontics and semantic neutrality. *Pacific Philosophical Quarterly*, <https://doi.org/10.1111/papq.12046>.
- Cariani, F. (2016). Deontic modals and probabilities: One theory to rule them all? In N. Charlow & M. Chrisman (Eds.), *Deontic modality* (pp. 11–46). Oxford: Oxford University Press.
- Cariani, F., Kaufmann, M., & Kaufmann, S. (2013). Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy*, 36, 225–259. <https://doi.org/10.1007/s10988-013-9134-4>.

- Carr, J. (2012). Deontic modals without decision theory. In E. Chemla, V. Homer, & G. Winterstein (Eds.), *Proceedings of Sinn und Bedeutung* (Vol. 17, pp. 167–182). Paris: ENS Paris.
- Carr, J. (2015). Subjective ought. *Ergo*, 2, 678–710. <https://doi.org/10.3998/ergo.12405314.0002.027>.
- Charlow, N. (2013a). Conditional preferences and practical conditionals. *Linguistics and Philosophy*, 36, 463–511. <https://doi.org/10.1007/s10988-013-9143-3>.
- Charlow, N. (2013b). What we know and what to do. *Synthese*, 190, 2291–2323. <https://doi.org/10.1007/s11229-011-9974-9>.
- Charlow, N. (2014a). Logic and semantics for imperatives. *Journal of Philosophical Logic*, 43, 617–664. <https://doi.org/10.1007/s10992-013-9284-4>.
- Charlow, N. (2014b). The problem with the Frege–Geach problem. *Philosophical Studies*, 167, 635–665. <https://doi.org/10.1007/s11098-013-0119-5>.
- Charlow, N. (2015). Prospects for an expressivist theory of meaning. *Philosophers' Imprint*, 15, 1–43.
- Charlow, N. (2016). Decision theory: Yes! Truth conditions: No!. In N. Charlow & M. Chrisman (Eds.), *Deontic modality* (pp. 47–81). Oxford: Oxford University Press.
- Chrisman, M. (2012). On the meaning of 'ought'. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 7, pp. 304–332). Oxford: Oxford University Press.
- Dowell, J. L. (2012). Contextualist solutions to three puzzles about practical conditionals. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 7, pp. 271–303). Oxford: Oxford University Press.
- Dreier, J. (1999). Transforming expressivism. *Noûs*, 33, 558–572. <https://doi.org/10.1111/0029-4624.00193>.
- Gibbard, A. (1990). *Wise choices, apt feelings*. Cambridge: Harvard University Press.
- Gibbard, A. (2003). *Thinking how to live*. Cambridge: Harvard University Press.
- Gibbard, A., & Harper, W. L. (1981). Counterfactuals and two kinds of expected utility. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Conditionals, belief, decision, chance and time* (pp. 153–190). Dordrecht: Springer.
- Hawthorne, J. & Stanley, J. (2008). Knowledge and action. *The Journal of Philosophy*, 105, 571–590. <http://www.journalofphilosophy.org/articles/issues/105/10/1.pdf>.
- Joyce, J. (1999). *The foundations of causal decision theory*. Cambridge: Cambridge University Press.
- Kahneman, D. (2003). A psychological perspective on economics. *The American Economic Review*, 93, 162–168. <https://doi.org/10.1257/000282803321946985>.
- Katz, G., Portner, P. & Rubinstein, A. (2012). Ordering combination for modal comparison. In A. Chereches (Ed.), *Proceedings of SALT* (Vol. 22, pp. 488–507). CLC Publications. <https://doi.org/10.3765/salt.v22i0.2647>.
- King, J. C. (2014). The metasemantics of contextual sensitivity. In A. Burgess & B. Sherman (Eds.), *Metasemantics: New essays on the foundations of meaning* (pp. 97–118). Oxford: Oxford University Press.
- Kolodny, N., & MacFarlane, J. (2010). Ifs and oughts. *Journal of Philosophy*, 107, 115–143. <https://doi.org/10.5840/jphil2010107310>.
- Kratzer, A. (1977). What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, 1, 337–355. <https://doi.org/10.1007/BF00353453>.
- Kratzer, A. (1981). The notional category of modality. In H. Eikmeyer & H. Rieser (Eds.), *Words, worlds, and contexts* (pp. 38–74). Berlin: De Gruyter.
- Kratzer, A. (1991). Modality. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 639–651). Berlin: De Gruyter.
- Kratzer, A. (2012). *Modals and conditionals*. Oxford: Oxford University Press.
- Lassiter, D. (2011). *Measurement and modality*, Ph.D. Dissertation. NYU.
- Lassiter, D. (2016). Linguistic and philosophical considerations on Bayesian semantics. In N. Charlow & M. Chrisman (Eds.), *Deontic modality*. New York: Oxford University Press.
- MacFarlane, J. (2011). Epistemic modals are assessment-sensitive. In B. Weatherston & A. Egan (Eds.), *Epistemic modality* (pp. 144–178). Oxford: Oxford University Press.
- MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. Oxford: Oxford University Press.
- MacFarlane, J. (2016). Vagueness as Indecision. *Aristotelian Society Supplementary Volume*, XC, 255–283. <https://doi.org/10.1093/arisup/akw013>.
- Maitra, I., & Weatherston, B. (2010). Assertion, knowledge, and action. *Philosophical Studies*, 149, 99–118. <https://doi.org/10.1007/s11098-010-9542-z>.

- Neta, R. (2009). Treating something as a reason for action. *Noûs*, 43, 684–699. <https://doi.org/10.1111/j.1468-0068.2009.00724.x>.
- Plunkett, D. & Sundell, T. (2013). Disagreement and the semantics of normative and evaluative terms. *Philosopher's Imprint*, 13, 1–37. <http://hdl.handle.net/2027/spo.3521354.0013.023>.
- Portner, P. (2007). Imperatives and modals. *Natural Language Semantics*, 15, 351–383. <https://doi.org/10.1007/s11050-007-9022-y>.
- Portner, P. (forthcoming). Commitment to priorities. In D. Fogal, D. Harris, & M. Moss (Eds.), *New work on speech acts*. Oxford: Oxford University Press.
- Rawls, J. (1971). *A theory of justice*. New York: Oxford University Press.
- Ridge, M. (2014). *Impassioned belief*. Oxford: Oxford University Press.
- Savage, L. J. (1972). *The foundations of statistics*. New York: Dover.
- Schroeder, M. (2008a). *Being for*. Oxford: Oxford University Press.
- Schroeder, M. (2008b). Expression for expressivists. *Philosophy and Phenomenological Research*, 76, 86–116. <https://doi.org/10.1111/j.1933-1592.2007.00116.x>.
- Schroeder, M. (2008c). How expressivists can and should solve their problem with negation. *Noûs*, 42, 573–599. <https://doi.org/10.1111/j.1468-0068.2008.00693.x>.
- Schroeder, M. (2011). Ought, agents, and actions. *The Philosophical Review*, 120, 1–41. <https://doi.org/10.1215/00318108-2010-017>.
- Silk, A. (2013). Truth-conditions and the meanings of ethical terms. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 8, pp. 195–222). Oxford: Oxford University Press.
- Silk, A. (2014). Evidence sensitivity in weak necessity deontic modals. *Journal of Philosophical Logic*, 43, 691–723. <https://doi.org/10.1007/s10992-013-9286-2>.
- Silk, A. (2015). How to be an ethical expressivist. *Philosophy and Phenomenological Research*, 91, 47–81. <https://doi.org/10.1111/phpr.12138>.
- Stalnaker, R. (2011). *Mere possibilities: Metaphysical foundations of modal semantics*. Princeton: Princeton University Press.
- Steele, K. (2007). Distinguishing indeterminate belief from “risk-averse” preferences. *Synthese*, 158, 189–205. <https://doi.org/10.1007/s11229-006-9119-8>.
- Swanson, E. (2008). Modality in language. *Philosophy Compass*, 3, 1193–1207. <https://doi.org/10.1111/j.1747-9991.2008.00177.x>.
- von Fintel, K. (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.
- von Fintel, K., & Gillies, A. S. (2011). Might made right. In A. Egan & B. Weatherson (Eds.), *Epistemic modality* (pp. 108–130). Oxford: Oxford University Press.
- Weiner, M. (2005). Must we know what we say? *The Philosophical Review*, 114, 227–251. <https://doi.org/10.1215/00318108-114-2-227>.
- Weirich, P. (1980). Conditional utility and its place in decision theory. *The Journal of Philosophy*, 77, 702–715. <https://doi.org/10.5840/jphil1980771110>.
- Willer, M. (2012). A remark on iffy oughts. *Journal of Philosophy*, 109, 449–461. <https://doi.org/10.5840/jphil2012109719>.
- Willer, M. (2013). Dynamics of epistemic modality. *The Philosophical Review*, 122, 45–92. <https://doi.org/10.1215/00318108-1728714>.
- Yalcin, S. (2010). Probability operators. *Philosophy Compass*, 5, 916–937. <https://doi.org/10.1111/j.1747-9991.2010.00360.x>.
- Yalcin, S. (2012a). Bayesian expressivism. In *Proceedings of the Aristotelian Society* CXII, Part 2 (pp. 123–160). <https://doi.org/10.1111/j.1467-9264.2012.00329.x>.
- Yalcin, S. (2012b). Context probabilism. In M. Aloni, V. Kimmelman, F. Roelofsen, G. Sassoon, K. Schulz, & M. Westera (Eds.), *Proceedings of the 18th Amsterdam Colloquium* (pp. 12–21). ILLC. https://doi.org/10.1007/978-3-642-31482-7_2.
- Yalcin, S. (2014). Semantics and metasemantics in the context of generative grammar. In A. Burgess & B. Sherman (Eds.), *Metasemantics: New essays on the foundations of meaning* (pp. 17–54). Oxford: Oxford University Press.