

Scuola Dottorale di Ateneo

Graduate School

Dottorato di Ricerca in Filosofia

Ciclo 27

Anno di discussione 2016

The Meaning and Use of Normative Language

Tesi di dottorato di Giulia Pravato

Matricola: 825720

Coordinatore del dottorato

Prof. Carlo Natali

Tutore del dottorando

Prof. Luigi Perissinotto

Abstract

This work explores a new hypothesis about the semantics and pragmatics of normative language. The core conjecture is that an utterance of a normative sentence (like “Harry ought to keep his promise”) expresses a plurality of (classical, bivalent) propositions, some of which are true and some of which are false. This picture is inspired by an analogy with a certain view about the semantics of vague language. On this view, vague sentences (like “Harry is bald”) are associated with a multitude of candidate propositions, one for each admissible precisification of the vague word. However, as I attempt to show, there is a crucial difference between vague language and normative language. Some vague sentences are such that they are true on all their admissible precisifications. We can safely ignore their multiplicity of meanings, as it were. Normative sentences aren’t like that. For any action-type, you may find a normative standard which counts that action into the extension of “wrong” and one that does not, and yet such that both respect all the relevant meaning-fixing facts (use, intentions, conventions, etc.). Normative language, we could say, is *wildly* or *irresolvably vague*. In a slogan: all cases are borderline.

In my view, this hypothesis can be accepted by normative absolutists (or invariabilists, as I call them) and by normative relativists (variabilists, in my terminology) alike. What distinguishes the two is that the latter think that the competing normative standards are equally correct from a *metaphysical* point of view, and not just from the point of view of semantic/conceptual competence. Ultimately, this package of maximal vagueness and metaphysical relativism, which I call *Trivial Descriptivism*, is the overall approach defended in this essay.

The dissertation consists of three chapters. In Chapter 1, I contrast two metanorma-

tive approaches that I call *invariabilism* and *variabilism*. Invariabilism and variabilism are views about the *metaphysics* of normativity. The first is the view that there is a single uniquely correct normative standard (or perhaps a small set of such standards); the second is the view that there are many equally acceptable normative standards. Besides setting the stage, the main goal of this chapter is to reject invariabilism. The main argument proceeds by highlighting an asymmetry in our ascriptions of semantic competence/incompetence to users of descriptive *vs* normative language, and by suggesting that this asymmetry is best explained by a variabilist approach.

Chapter 2 surveys the extant semantic options for variabilists, and their respective virtues and vices. I discuss four proposals—contextualism, expressivism, truth-relativism, and the error theory—and I raise some (more or less original) worries for each of them. Some special attention is devoted to the truth-relativist framework, the newest and less explored of the four approaches, at least in the metaethical/metanormative literature.

Chapter 3 articulates, explores and motivates my own positive proposal, Trivial Descriptivism. The main insight, as summarized above, is framed by adopting a “cloudy semantics”: the envisaged multiplicity of meanings is built into the semantics itself. More precisely, the semantic content of a normative sentence in context is a set or cloud of propositions—one for each possible moral standard—rather than a single proposition. I then turn to explain how the “cloudy semantic facts” bear on explaining how the target sentences are conventionally used in conversation, and relatedly on predicting the various norms governing the production, acceptance, rejection, and retraction of normative claims in the dynamics of typical conversations. Perhaps unsurprisingly, I propose that it is the subjects’ normative outlooks, in particular the truth/falsity of the propositions in the cloud that correspond to the standards they accept, that play a role in determining the correctness/incorrectness of the various attitudes and speech-acts.

Acknowledgments

My biggest thanks must go to my main supervisor, Max Kölbl.

Contents

Abstract	1
Acknowledgements	3
1 Variabilism and Invariabilism	6
1.1 Introduction	6
1.2 Invariabilism and Variabilism Defined	10
1.3 An Argument Against Invariabilism	12
1.3.1 Denying the Asymmetry: Normative Conceptual Truths	17
1.3.2 Extensional Constraints* <i>Sans</i> Paradigm Cases	19
1.3.3 A Different Suggestion	24
1.3.4 Conclusion	26
1.4 Variability and Reducibility	27
1.5 Natural Language Semantics and Metanormative Inquiry	28
2 What Normative Words Mean: Some Old and New Stories	34
2.1 Introduction	34
2.2 Contextualism	35
2.2.1 Introducing Contextualism	35
2.2.2 Virtues and Vices	37
2.3 Expressivism	47
2.3.1 Introducing Expressivism	47
2.3.2 Virtues and Vices	49
2.4 Relativism	53

2.4.1	Introducing Relativism	53
2.4.2	Virtues and Vices	53
2.5	The Error Theory	69
2.5.1	Introducing the Error Theory	69
2.5.2	Virtues and Vices	71
3	Trivial Descriptivism	74
3.1	Introduction	74
3.2	Cloudy Semantics Explained	78
3.2.1	Multiplicity and Indeterminacy	79
3.2.2	A Toy Cloudy Semantics	82
3.2.3	Some Worries and Replies	85
3.3	Pragmatics: A General Framework	88
3.3.1	Contexts as Scoreboards	89
3.3.2	Assertions as Proposals	90
3.3.3	Monadic Truth	94
3.4	Trivial Descriptivism and the Explanatory Challenges	96
3.4.1	Assertion	97
3.4.2	Agreement and Disagreement	102
3.4.3	Retraction	105
3.5	Comparison With Similar Approaches	106
3.6	Metasemantic Issues	113
3.7	Keeping the Cloud Out Of The Semantics	116
	Bibliography	117

Chapter 1

Variabilism and Invariabilism

1.1 Introduction

Invariabilists about thin normative statements—roughly, statements involving terms like “good”, “right”, “ought”, “reason”, etc.—believe that there is a single uniquely correct normative standard, or perhaps a small set of such standards. What is a standard? Roughly, we can think of it as something that determines a complete and consistent assignment of truth-values to normative sentences. In other words, a standard doesn’t leave any case underdetermined or overdetermined.¹

Consequently, invariabilists believe that if A uses a thin normative sentence φ —i.e. a sentence like “ x is good”, “ x is right”, “you ought to do x ”, “you have most reason to do x ”²—to make a sincere assertion and B uses its negation $\neg\varphi$ to make a sincere assertion, then in all or most cases one of them is at fault and the other is not. I should qualify this

¹This is modelled on Gibbard (1990) and (2003a), which idealizes over inconsistencies and uncertainties in his characterization of a system of norms (1990) or hyperplan (2003). I will bracket other interesting questions—e.g. whether we should think of a standard as a set of *propositions* describing an ideal world, or rather as a set of (conditional) *imperatives*. See Boghossian (2006) and Field (2009) for discussion.

²Two *caveats*. First: I’m using “normative” as a catch-all term for both *deontic* words (like “right” and “wrong”) and *evaluative* ones (like “good” and “bad”). Thin terms are contrasted with other broadly normative terms, for example so-called thick terms like “courageous” or “rude”. I will not attempt to give an informative characterization of the thin/thick distinction. The literature on the topic shows that any definition is likely to be theory-laden and unacceptable to this or that theorist. Second: for simplicity, I shall talk of normative *terms*, even if it would be more appropriate to talk of normative *uses/readings* of some terms. Moreover, to keep the discussion tractable, I’ll mostly focus on adjectives (like “wrong”) and modals (like “ought”) and set aside terms belonging to other syntactic categories, as much as more-fined grained distinctions one can draw in the above categories (e.g. that between so-called weak necessity modals like “should”, and strong necessity modals like “must”).

claim. An invariabilist is happy to grant that two sincere utterances of “*I* ought to keep *this* promise” and “It is not the case that *I* ought to keep *this* promise” performed by A and B respectively, may not be such that either A or B is at fault. So, the invariabilist thesis is plausibly restricted to those normative sentences whose non-normative components are not context-sensitive. Alternatively, the thesis can be restricted to certain uses of normative sentences, namely those where all the contextual parameters, except the one for a normative standard (if required by one’s semantics), have the same values.

With this in mind, we can say that according to invariabilists, normative statements are pretty much like geographical statements: if A sincerely and assertorically utters “Los Angeles is in California” and B sincerely and assertorically utters “Los Angeles is not in California”, one of them is at fault and the other is not, and this turns on what the world is geographically like.

Normative *variabilists* think that there are many, actual or possible, correct normative standards. If you’re a variabilist, you may be an *error theorist* and think that if A sincerely and assertorically utters “eating fish is wrong” and B sincerely and assertorically utters “eating fish is not wrong”, both A and B must be at fault—their utterances are equally untrue. On one way of construing this view, normative statements are like statements involving empty terms, such as “ghost”, “witch” or “phlogiston”.

Alternatively, you may be a *success theorist* and think that neither A nor B may be at fault—both may be speaking the truth (or at least both may be speaking correctly). There are various ways of making the “joint truth” (or “joint correctness”) hypothesis coherent. A natural one is to look at the model of *indexicals*, such as “I”, “now”, “today”, etc. If A says “I’m American” and B says “I’m not American”, they may both be speaking the truth; the word “I” denotes A in the mouth of A and B in the mouth of B, after all. Similarly, if “wrong” denotes two distinct properties when used by A and B—i.e. *wrong according to A’s body of norms* and *wrong according to B’s body of norms*, respectively—it is perfectly clear why two syntactically contradictory statements such as “eating fish is wrong” and “eating fish is not wrong” can both be true. (There are, of course, other ways of being “joint truth” theorists, as we shall see in a moment).

To wrap up, if A sincerely and assertorically utters “eating fish is wrong” and B sincerely and assertorically utters “eating fish is not wrong”, we can say one of the following three things: (i) one and only one of the two utterances is true; (ii) the two utterances are equally untrue; (iii) the two utterances may both be true (or at least correct). (i) is called “invariabilism”, (ii) and (iii) “variabilism”.

Intuitively, and minimally, the truth-value (and/or the correctness/incorrectness) of a given declarative utterance depends on the conventional meaning of the sentence uttered and on whether the conditions or constraints imposed by such meaning are satisfied or not. I think that one can accept this claim independently of one’s semantic-theoretic preferences, e.g. regardless of whether one opts for a truth-theoretic analysis (meaning is *truth-conditions*) or for a use-theoretic one (meaning is *use-* or *assertability-conditions*).

That said, I think there are good reasons for discarding (i). The main argument of this chapter, which I shall be developing in §1.3, follows the following rough outline: (i) could only be true if normative sentences meant such-and-such; but they don’t. So (i) is false.

The live options are then (ii) and (iii). There are four main views one might adopt, if (i) is off the table: Contextualism, Relativism, Expressivism and the Error Theory. Oversimplifying a bit, *Contextualists* believe that a normative sentence φ can express different propositions in different contexts of utterance, for instance when used by different speakers. *Relativists* believe that φ expresses the same proposition across contexts, but that this proposition may be true for an assessor and false for another. *Expressivists* think that φ doesn’t purport to express a proposition (at least in any sense in which propositions have an important explanatory leverage); instead, it expresses some non-cognitive attitude and/or some imperative. *Error theorists* believe that φ either fails to express a proposition while it purports to express one, or it expresses a proposition that is false or neither true nor false. Either way, φ turns out to be untrue.

The general goal of this thesis is to evaluate the prospects for a (largely) novel variabilist semantics of normative sentences. The main idea is that a normative sentence φ (in context) expresses a plurality of (classical and bivalent) propositions, some of which

are true and some of which are false, instead of a single proposition (perhaps with relative truth-values) or no proposition at all. I will develop my proposal in Chapter 3, after discussing the alternative existing approaches in Chapter 2.

So far, my examples have only concerned a subset of the normative occurrences of “good”, “right”, and “ought”, namely their *moral* uses, as displayed in sentences like: “one ought to keep one’s promises”, “it is right to help the poor”, “cutting taxes is good”. Assertions of these sentences are generally called *categorical*, and contrasted with *instrumental* or *hypothetical* ones, insofar as the relevant obligations and values invoked by the former are not, on the face of it, conditional on having some (implicitly or explicitly specified) end or goal (as in: “[if you want to loose weight] you ought to start a diet) or relativized to some particular system of rules (as legal obligations are, for example).

But there are other normative statements, besides moral ones, that arguably possess the relevant categoricity—for example, claims involving *epistemic* norms (i.e. norms about what one ought to believe) and *aesthetic* norms (i.e. norms about what one ought to feel, like, enjoy, etc.). I think that most of the remarks I’ll be making throughout this essay equally apply to aesthetic and epistemic normativity. To keep the discussion tractable, though, I’ll mainly focus on moral statements.

Giving an informative characterization of what evaluations count as moral is not an easy task, but I think we have an intuitive grasp of what makes a claim moral, as opposed to epistemic, or aesthetic. There are some paradigm examples of moral claims—“stealing is wrong”, “one ought to help the poor”, “eating meat is bad”—and some paradigm examples of non-moral and yet normative claims—take “one ought not to believe p and $\neg p$ ” or “Scorsese is a great director”, an epistemic and aesthetic statement, respectively. So I will leave the subject here.³

This chapter is structured as follows. In §1.2 I fine-tune my characterization of

³One may think that *moral* standards are concerned with the assessment of actions or of action-related subjects. But one may evaluate actions also with respect to traffic rules. Phillips (1997) helpfully individuates four general characterizations of the concept of morality, i.e. of what makes an action guide a morality: (i) the fact that “it plays a certain role within the deliberative life of an individual”; (ii) the fact that it has a certain sort of content, e.g. being related to “human good and harm”, as Foot (1978) suggests; (iii) the fact that it plays a certain *social* role, like solving coordination problems and facilitating human cooperation; (iv) the fact that it is connected to the right sorts of emotions, e.g. guilt, resentment, and approval.

invariabilism and variabilism. In §1.3 and §1.4 I present my case against invariabilism. §1.5 introduces a semantic framework that will be used throughout this work.

1.2 Invariabilism and Variabilism Defined

In the *Introduction* I defined “invariabilism” as the view according to which there is a single uniquely correct normative standard or perhaps a small set thereof, in virtue of which normative statements are true or false.⁴ With an important qualification, I also said that this view entails that if A asserts a normative sentence φ and B asserts its negation $\neg\varphi$, then one and only one of them speaks the truth.

This characterization is broad and liberal enough to include views on which normative properties are reducible to/constituted by natural properties (*à la* Jackson (1998), or *à la* Brink (1989) respectively)⁵, views on which normative properties are not only reducible to/constituted by natural properties, but also mind-dependent (as in non-Humean versions of constructivism, *à la* Korsgaard (1996) or *à la* Smith (1994)), and perhaps minimalist expressivists (*à la* Blackburn (1993) or Gibbard (1990) and (2003a)) who help themselves to deflated conceptions of “fact”, “representation”, and related notions.

My aim in working with this distinction is to carve up the terrain in a way which fits my purposes—specifically, to distinguish between those who affirm and those who

⁴Harman (2015) recently used the term “moral relativism” to capture something like the negation of the latter. He defines moral relativism as the thesis that “there is not a single true morality”. I think the word “invariabilism” is less fraught. “Relativism” has recently become associated with a view about the semantics of moral statements (Harman himself previously used the term to refer to a logico-semantic claim, in his famous (1975)). Obviously, this is just a worry about labelling.

⁵The *reductive naturalist* (see, for instance, Jackson (1998)) thinks that a normative property, by which I mean a property that can be ascribed with a normative predicate, is identical to a property picked out by a descriptive, i.e. non-normative, predicate or by a conjunction/disjunction of descriptive predicates. For instance, the property ascribed by “morally right” or “morally good” is identical to a property ascribed by a psychological, sociological, or biological predicate (or any other predicate drawn from the vocabulary of some science) like “being a happiness-maximizer”, “being a welfare-maximizer”, “being a healthiness-maximizer”.

The *non-reductive naturalist* (see, for instance, Brink (1989)) thinks that the properties ascribed with normative predicates are not identical to the properties picked out by descriptive predicates (or conjunctions/disjunctions thereof). Rather, they are *constituted*, *composed*, or *realized* by natural properties. The relationship between normative and non-normative properties would be analogous to that between mental and physical properties, according to some non-reductivist and yet physicalist views of the mental. The upshot is that normative properties are *metaphysically continuous* with other natural properties, in virtue of being empirically discoverable and causally efficacious, even if they don’t (or don’t need to) have characterizations in non-normative terms.

deny the existence of a single uniquely correct or adequate normative system, standard, or perspective.

Some people would use the term “realism” to name what I’m naming “invariabilism”. But most definitions of “realism” rule out constructivist and expressivist views from the realist camp.⁶ Moreover, some authors want to count contextualism and relativism as realist views, on the grounds that they admit the existence of normative facts, they simply take them to be relational facts.⁷ To avoid terminological distractions, thus, I prefer to use a less fraught dichotomy.

It is often implicitly assumed that any metanormative view must be held uniformly for all normative sentences: more specifically, it is assumed that if one thinks that *some* normative sentences φ_i are such that either φ_i are absolutely true (i.e. true regardless of who utters or assesses it) or their negations $\neg\varphi_i$ are absolutely true, one will thereby think that the same is true of *all* normative sentences. This seems a mistake. Some (liberal) invariabilist may grant that there are some normative sentences regarding which there is no (context-independent) fact of the matter as to whether they are true or false.⁸ The possibility of this sort of liberal invariabilism motivates the disjunctive formulation of the view I’ve given above (invariabilism was there defined as the view that there is a single, *or a small set of*, correct normative standards).

By the same token, a variabilist worth the name may want to make some exceptions for sentences like “torturing babies for fun is wrong” and its ilk. To be clear: I think that the correct view is a sweeping variabilism, which holds for any atomic and substantive normative sentence. But the tempered variabilism just outlined looks like a coherent view, and as such should be included in the logical space of possibilities.

⁶Cf., for instance, Shafer-Landau’s definition of realism: “Realists believe that there are moral truths that obtain independently of any preferred perspective, in the sense that *the moral standards that fix the moral facts are not made true by virtue of their ratification from within any given actual or hypothetical perspective*. That a person takes a particular attitude toward a putative moral standard is not what makes that standard correct” (Shafer-Landau (2003), p. 15).

⁷Cf. Kalderon (2005), pp. 213-214. Harman (2015) also suggests that moral relativism, defined as “the view that there is not a single true morality” is compatible with realism, for is compatible with the view that “the relevant relations are real” in roughly the same way in which the relation of being at rest (or moving) relative to a spatio-temporal frame of reference is real, or the relation of being legal relative to a legal code is real.

⁸See, for instance, Shafer-Landau (1994) and (1995) according to which some situations may be morally indeterminate.

1.3 An Argument Against Invariabilism

Consider ordinary descriptive terms like “chair”, “mountain”, “bald”, “rich” and “red”. One striking feature of these terms is that they present some *paradigm cases* of application.

Let’s define a paradigm case as follows:

- (1) **x is a paradigm case of the use of ‘F’** =_{df} x is such that if someone in apparently good epistemic conditions refuses to apply ‘F’ to x , then there is an initial temptation to regard the subject in question as linguistically incompetent.

The notion of “good epistemic conditions”, and consequently of “apparently good epistemic conditions” is notoriously tricky to spell out. Intuitively, we can think of it as a matter of being well-placed to find out the truth about a certain domain. This allows for variations across domains, and only requires actual evidence instead of full evidence or other more idealized notions; similarly, it doesn’t require superhuman cognitive powers. I use the term “apparently” because what matters is what the person who ascribes competence/incompetence *believes*. So here is the final, admittedly pretty vague, formulation:

- (2) **A is in apparently good epistemic conditions** =_{df} A seems to possess the evidence that is relevant to the application of ‘F’ and seems free from failures of rationality or other epistemic vices.

Imagine someone who is informed about the number and distribution of Vin Diesel’s cranial hairs and yet says: “Vin Diesel is not bald”. All else being equal, it is fairly plausible to infer that the subject doesn’t know what “bald” means. Similar considerations apply to someone who denies that a 14 year-old boy is young, that Bill Gates is rich while knowing its income, or that the Everest is a mountain, while being informed about the geographical facts. *Ceteris paribus*, there is a strong initial temptation to regard the person in question as linguistically incompetent. Our *default attitude*, that is, is to think that the speaker must be incompetent or confused about the meaning of ‘F’.

Two things are especially important to highlight. First, I’m *not* saying that if an object, event, fact, etc. a is a paradigm case of application of ‘F’ (or, for short, a

paradigmatic F) then it is *correct* to apply ‘F’ to *a*, and consequently *a is F*.⁹ This is why the italicized hedges are important. The initial ascription of incompetence is compatible with revising our judgment if further information comes along. Suppose we learn that the person denying that Vin Diesel is bald is a philosophy student who after being exposed to the sorites paradox concludes on the basis of that reasoning that Vin Diesel isn’t bald, for someone who has 400.000 cranial hairs isn’t bald, and if someone who has *n* cranial hairs isn’t bald, then someone who has (*n*–1) cranial hairs isn’t either, etc. You might think that he’s confused, but deny that he displays semantic incompetence. In light of this, I think that the claim that descriptive terms display paradigm cases of application in the sense defined above is acceptable regardless of one’s metasemantic preferences. Metasemantic considerations become relevant when we move from the claim that *a* is *paradigmatically* called ‘F’ to the claim that *a* is *correctly* called ‘F’. Like I said, it’s only the first claim that matters to my current purposes.

Second, paradigm cases of F-ness are not the same thing as what we may call *Moorean cases* of F-ness. Say that a Moorean case of F-ness is an *x* such that the vast majority of users of ‘F’ *believe* it to be in the extension of ‘F’. The notion of a Moorean case is an *epistemic* notion, while that of a paradigm case is a semantic notion. Perhaps everyone believes that the world will not come to an end tomorrow, but this is needn’t be constitutive of semantic competence with any of the words involved.

Let us look at normative predicates. Are there paradigm cases of application of ‘F’, when F is a normative predicate? On the face of it, the answer seems negative. Take your best candidate for a paradigm case of wrongness, for instance the action-type of *torturing babies for fun*. It seems that someone can know the exact distribution of the morally relevant natural properties of that action-type—like that it produces far more pain than pleasure, say—and yet sincerely and competently deny that that action has the property of being wrong.¹⁰

⁹I say “consequently” because I’m assuming that the following principle holds: *x* is F iff *x* is correctly called ‘F’.

¹⁰Some philosophers disagree. I will return to this below.

More generally, we do not automatically view normative dissenters or outsiders as linguistically incompetent. We do not think that the cannibal, the slaveholder, the sociopath, the futurist and the iconoclast are misusing language when, despite being fully informed about the relevant non-normative facts, they sincerely utter sentences like “cannibalism is permissible”, “slavery is just”, “torturing babies for fun is permissible”, “Venice is not a beautiful city”, or “we have all along been acting contrary to reason”. At the very least, this is not our initial sense or default reaction. *Prima facie* (or *ceteris paribus*), normative outsiders can say what they say without violating any linguistic convention.¹¹

The hypothesis is that this holds for any, or almost any, normative sentence that is atomic—i.e. of the form x is F (or A ought to ϕ , etc.)—at least inasmuch as this sentence is used to make a categorical claim. It is important that the sentence be *atomic*. Surely, I can’t be competent and deny compound sentences like “wrongful killing is wrong”, “if something is wrong, then it is wrong”, or belief-ascriptions like “Mia believes that eating fish is wrong” if Mia really believes that.

Why *almost* any atomic sentence? There are some exceptions. (i) First, there may be (uninteresting) analytic connections like: “murder is wrong”. (ii) Second, normative predicates (or at least some of them) have a domain of application which seems restricted to a certain class of items (moral terms, for example seem to apply to, e.g., actions, intentions, policies, etc.). For instance, it seems that “is morally wrong” neither applies nor fails to apply to, say, the Pacific Ocean, a prime number, or a table. Its domain of application seems restricted to items that are suitably related to actions, intentions, etc. If you think that alleged “category mistakes” are simply cases of falsehoods, you can just say that those sentences are false. For our purposes, nothing turns on this. It is actually controversial whether we should build syntactical constraints like those into the meaning of normative terms. We can perhaps imagine someone who, based on some bizarre metaphysical theory comes to think that the Pacific Ocean is morally wrong (or right). He may be massively confused, but needn’t display incompetence. Surely, though,

¹¹Street (2009) presents a number of philosophically popular “ideally coherent eccentrics”, ranging from Hume’s pure egoist who “prefers the destruction of the whole world to the scratching of [his] finger”, to Parfit’s Future-Tuesday-Indifferent “who always prefers agony on a future Tuesday to the slightest pain on any other day”.

that would be a *prima facie* evidence of semantic confusion. (iii) Third, even restricting our attention to actions, we can perhaps imagine further extensional constraints. Suppose someone claims that it was morally wrong for the ancient Egyptians to be heliocentrists, or that it is morally wrong to breath, or to fall ill. Again, we would be initially tempted to think that the person doesn't master the meaning of "morally wrong" (for instance, in the first case, she confuses moral wrongness with a different kind of wrongness). In general, not all descriptive features seem normatively relevant, at least given the kind of creatures we are. *Intentional* and *phenomenological* facts—e.g. the intentions guiding an action and the pleasure/pain produced by the action—score definitely higher than, say, *meteorological* and *temporal* facts—e.g. the weather conditions and the day of the week on which an action is performed.

Do (i), (ii), and (iii) prove that there are paradigmatic uses of 'F' in the relevant sense? No. While (i) only trivially meets the condition, (ii) and (iii) exemplify cases where the term *paradigmatically fails* to apply.

If what I have said so far is on the right track, we seem to observe an important asymmetry in our ascriptions of semantic competence/incompetence:

Asymmetry: While there are many paradigmatic uses/applications of 'F' when F is an ordinary descriptive term, there seems to be no such uses when 'F' is a normative term.

What's the importance of **Asymmetry**, and why should it help us build a case against the invariabilist? I want to put forward the following strategy. Ordinary descriptive terms provide the best candidates for terms that comply with following condition:

Extensional Constraint: There are some *x*s (objects, individuals, events, etc.) such that it is correct for everyone to apply 'F' to *x*.

Do descriptive terms really meet **Extensional Constraint**? This is controversial. Philosophers like Unger (1979), Braun and Sider (2007), and Beall (2010) claim that, for reasons having essentially to do with the sorites paradox, vague predicates like "bald" do not apply to anybody. More generally, there is almost no descriptive predicate that hasn't been the target of arguments purporting to show that the predicate is empty. Of

course, one might argue that these philosophers are mistaken, and that **Extensional Constraint** is true of vague predicates. The quickest way to do so would be to say that if x is paradigmatically called ‘F’, then x is correctly called F. But as already mentioned above, this inference is problematic. Moreover, and more importantly, we don’t need to assume that much. The following revised constraint will do for our purposes:

Extensional Constraint*: There are some xs (objects, individuals, events, etc.) such that either it is correct for everyone to apply ‘F’ to x or it is mistaken for everyone to apply ‘F’ to x .

It seems to me that **Extensional Constraint*** holds unconditionally for descriptive terms like “chair”, “bald” and its ilk, and that this is common ground between metanormative variabilists and invariabilists. This suggests the following argument:

(P1): For many expressions ε that ε uncontroversially satisfy **Extensional Constraint***, there seems to be many paradigm cases of applications of ‘ ε ’, i.e. many examples of applications of ‘ ε ’ such that denials of them, under good epistemic conditions, create an initial temptation to ascribe incompetence to the speaker.

(P2): Normative expressions present no such paradigm cases of application.

(C): Either normative expressions meet **Extensional Constraint*** without featuring paradigm cases, or normative expressions don’t satisfy **Extensional Constraint***.

The argument is admittedly modest, insofar as it is constituted by an existential (contrast: universal) premise, i.e. (P1), and it only delivers a disjunctive conclusion. In a way, this is how it should be: arguably, we cannot pull a metaphysical rabbit out of a (meta)semantic hat. Nonetheless, I think it raises a significant challenge for the invariabilist.

There are several things the latter could say in reply. First, he could deny (P1) by suggesting a view on competence that drops the very idea that ordinary terms feature paradigm cases of application. To me, this looks like a hopeless strategy. All the metase-mantic stories I’m aware of assume the existence of something like paradigm cases, in the

(very weak) sense defined above. I will thus set aside this option. Alternatively, he could deny the asymmetry from the other direction: he could reject **(P2)**, by saying that there are paradigm cases of application for normative terms. The next section, 1.3.1, is devoted to evaluating and discarding this possibility. Finally, the invariabilist could argue that the first disjunct of **(C)** is true. The extensions of normative predicates are fixed in a peculiar way, and this explains why **Extensional Constraint*** holds in spite of the lack of paradigmatic instances of F-ness. In §1.3.2 I examine this option, and show that it has some serious costs. Finally, in §1.3.3, I consider a question I left hanging in the air up to now, namely: what, if anything, counts as linguistic incompetence with normative terms, if paradigmatic applications are ruled out?

1.3.1 Denying the Asymmetry: Normative Conceptual Truths

Let us consider how one could object to the claim that competence with normative terms, and knowledge of the non-normative facts, is compatible with denying, say, that torturing babies for fun is wrong. There are, I think, two ways in which one could do that. First, one could argue that it is not the case that competent speakers' intuitions support the thesis. Second, one could claim that we shouldn't rely on competent speakers' intuitions in attributing semantic values to terms, for these intuitions are unreliable. I think the latter option is unconvincing.¹² The best bet for the invariabilist is to prove that the linguistic evidence doesn't support the view that normative outsiders can be competent users.

Balaguer (2011) presents the following as a “reasonably plausible” empirical hypothesis about the ordinary meaning of “wrong”:

It's built into our usage and intentions concerning the word 'wrong' that it is a property-ascribing term and that certain sorts of behaviours—e.g. killing babies just for money—are sufficient conditions for the correct applicability of this term (p. 384).

More recently, Cuneo and Shafer-Landau (2014) have claimed that some non-tautologous moral propositions are “nonnaturalistic conceptual truths” (note the difference with Bal-

¹²See above, §1.5.

aguer’s view, where nonnaturalism doesn’t play any role):

[...] there is a battery of substantive moral propositions, such as <that it is wrong to torture others just because they have inconvenienced you> and <that it is wrong to rape a child solely to indulge one’s lust>, that are also nonnaturalistic conceptual truths¹³.

I disagree, along with others (see below). We seem to have a clash of intuitions here. Who is right? I won’t insist on the correctness of my intuitions. Rather, I will suggest four points in reply.

(i) First, if these authors are correct, many theorists—variabilists or invariabilists—must fail to grasp the normative concepts. For these theorists typically don’t think that any of the supposed platitudes is a “paradigm case” (in the sense explicated above), even when they think that they are true. Although this is somewhat speculative, I think the numbers ultimately favor those who do not hear the relevant claims as semantically incoherent. In light of this, I’m more inclined to conclude that Balaguer’s and Cuneo and Shafer-Landau’s suggestions do not actually reflect competent speakers’ intuitions.

(ii) Second, seemingly competent speakers have divergent intuitions on the semantic correctness conditions of “wrong” and other normative terms. For instance, many authors seem to think that claiming that torturing babies for fun is not wrong does not involve making a *semantic* mistake (cf., among others, Schiffer (2003), Street (2009)). A possible approach could be to say that this clash of intuitions should be represented by our theory, for example by saying that our linguistic conventions don’t settle things one way or the other—i.e. it is indeterminate whether eccentric uses are linguistically incoherent or not. Although this would require a weakening of the original claim, it wouldn’t really affect the overall import of the foregoing considerations. Indeed, eccentric claims would still fail to be incoherent *simpliciter*.

(iii) Third, we could give a *pragmatic* explanation of Balaguer’s and other people’s intuition that the normative outsider misuses language. Compare, perhaps, pragmatic treatments of quantifier domain restriction, according to which the sentence “every bottle

¹³Conceptual truths are in turn defined as “propositions that are true in virtue of the essences of their constituent concepts”.

is empty” semantically expresses the false proposition *that every bottle in the universe is empty*, but pragmatically communicates the restricted proposition *that every bottle in the fridge is empty*.¹⁴ On this line, we could say that since *very often*, or in *normal conversations*, a standard that would verify the claim that torturing babies for fun is okay is not among the salient possibilities, we are inclined to draw the conclusion that it’s not a semantically acceptable interpretation. However, reflection on what we would say in some other possible scenarios suggests that the phenomenon is best treated as pragmatic in nature.

(iv) Fourth, let us grant for the sake of argument that there is a small set of sentences, denial of which manifests semantic incompetence. An example could be: “torturing babies for fun is wrong”. I don’t think this would make the problem for invariabilism go away, because such examples are exceptional and not typical of the moral norms that do play a role in our lives (for instance, by facilitating cooperation, planning, deliberation, and so on).

1.3.2 Extensional Constraints* *Sans* Paradigm Cases

One could argue that the first disjunct of the conclusion is true: normative terms comply with **Extensional Constraint*** despite lacking paradigm cases. How can this be? I can think of two options. Borrowing some terminology from Eklund (ms), the invariabilist could be “*metaphysically radical*” or “*metasemantically radical*”. On the first view, normative terms are special because their subject matter is somehow special: for instance, because the properties, relations, and facts picked out by normative predicates are intrinsically motivating and/or prescriptive, and in this sense utterly different from other properties we are familiar with. On the second view, normative terms are special because they “have their reference determined in a radically different way from that in which ordinary descriptive predicates get their reference determined” (ms, p. 16). Drawing from Wedgwood’s (2001) conceptual role semantics, Eklund (ms) fleshes out this option by saying that the reference of a normative term would be fully determined by its “norma-

¹⁴I’m just using this an analogy. I’m not committing myself to a pragmatic account of quantified domain restriction, in general.

tive role”—to a first approximation, its role in practical reasoning, deliberation, and so on.

Two important qualifications are due before we move on. First, being metasemantically radical doesn't necessarily involve being metaphysically radical, as both Wedgwood and Eklund stress. It may be that the question: “what is the goal of practical reasoning?” receives an answer that is fully compatible with naturalism. On the other hand, it is less clear how the converse entailment could fail. The invariabilist must posit a suitable reference-fixing mechanism to ensure that normative terms latch onto the desired entities. For otherwise, nothing guarantees that we are talking about rightness and wrongness, rather than about, say, rightness* and wrongness* or nothing at all, when we use normative language. And insofar as the invariabilist is a non-naturalist it is unclear how descriptivist or causal metasemantics could do the job. Second, invariabilism is only correct if normative predicates are not empty or radically indeterminate in reference. In this connection, Wedgwood (2001) claims that the reference of a normative term is determined by its role in practical reasoning and that it is a substantive metaphysical question, “what (if anything) is the goal of practical reasoning” (p. 19).

So, let's look at the metasemantically radical hypothesis. Wedgwood claims that to be semantically competent with a normative term like “x is (all things considered) a better thing than y for me to do at t” is to master the following rule: “Acceptance of 'B(x,y,me,t)' commits one to having a preference for doing x over doing y at time t” (N). In this respect, being competent with normative terms is analogous to being competent with the meaning of the logical constant (“and”, “or”, etc.) on the assumption that an inferentialist semantics for these terms is broadly correct. For instance, competence with the meaning of “and” involves the following dispositions: to infer *p and q* from a belief that p and a belief that q (I); and to infer p and to infer q from a belief that *p and q* (E).

Now, this is something that a non-cognitivist variabilist could be perfectly happy with. (I will return to this in the next section.). Importantly, Wedgwood's idea is that the rule (N) determines the reference/semantic value of the predicate in the same way in which (E) and (I) determine the reference of the connective. The idea is that in both

cases the semantic value is what makes all instances of the rules valid, i.e. it is “that semantic value that makes best sense of the fact that these rules are the basic rules” for the expressions in question (p. 10).

While this makes the account cognitivist, it doesn’t yet make it invariabilist. Why? Because the notion of *validity* is a purely *formal* notion. (I) doesn’t tell us when a belief is correct, or what it is for a belief to be correct; likewise (N) doesn’t tell us when a preference is correct, or what it is for a preference to be correct. Whether a belief is correct depends on what the goal or purpose of theoretical reasoning is, and whether a preference is correct depends on what the goal or purpose of practical reasoning is. As mentioned, Wedgwood notes that it is a substantive metaphysical question, “what (if anything) is the goal of practical reasoning” (p. 19). Further, Wedgwood’s account is not supposed to be a “non-circular conceptual analysis”, and as such is compatible with thinking that the *goal* or *purpose* of practical reasoning is itself a moral or anyway normative notion. Presumably, the invariabilist will say that there is a unique such goal and that a preference is correct only if the goal of practical reasoning permits one to have this preference.

This account might explain the possibility of EC* *sans* paradigm cases by highlighting a crucial disanalogy in the way in which the extensions of normative terms and those of ordinary terms are determined. In the former case, unlike in the latter, the inferential role fully determines the extension of the term. But is this sufficient to underwrite the kind of explanation we are after? It seems not.

First of all, insofar as we were looking for expressions that met EC* without displaying paradigm cases, terms which could then serve as models for normative terms, this is something that the analogy with logical connectives and theoretical reasoning simply doesn’t deliver. This is already problematic. If normative terms possess a fixed context-invariant extension, why should they be unlike all other predicates whose extensions are invariant in their failure to display paradigm cases of use? ¹⁵

¹⁵One could wonder whether the case of logic couldn’t be used in a different way, by arguing that moral rules are somehow like logical rules. Suppose that everyone in the debate of our concern grants that there are some logical rules such that either it is correct for everyone to accept them, or it is incorrect for everyone to accept them. In other words: there some rules such that faultless disagreements about

Second, regardless of that, there seems to be a wholly deflationary understanding of what counts as the goal of practical reasoning—say, to decide what to do. But, clearly, this would leave us with a simple non-cognitivist analysis. What the invariabilist needs is the existence of something like *the* goal of practical reasoning, and accordingly the existence of correct and incorrect preferences.

The problem I want to raise at this point can be put as a dilemma. Either there are paradigm cases of application of “the goal of practical reasoning”, or there aren’t. In light of the sort of linguistic intuitions canvassed above, I’m inclined to believe that there aren’t. This leads me to think that the only way in which the invariabilist can employ the idea of the goal of practical reasoning and talk of correct and incorrect preferences is by adopting some form of non-naturalism.

Non-naturalism comes in different stripes, but in general non-naturalists think that there are irreducible normative properties, facts, and relations, and that these entities are the truth-makers for normative judgments. Typically, they also claim that moral knowledge (or better, some moral knowledge, i.e. knowledge of the fundamental principles) is provided by intuition, a form of justification which is (i) non-inferential and which (ii) requires more than a “grasp of the meaning of *p*”. Non-naturalism faces a number of metaphysical and epistemological challenges. I will focus on what I think is one of its major costs, namely the fact that it posits inexplicable errors.

Consider, again, subjects with radically deviant normative outlooks. The invariabilist thinks that they are mistaken in virtue of what the irreducibly normative facts are. How can we explain their mistake in accordance with the intuitionist epistemology?

There are two options. First, one could say that the the outsider’s mistake can be ex-

the correctness of those rules is impossible. Call one such rule *R*. This assumption, call it *IC**, would play the same role as *EC** in the case of ordinary descriptive terms. One might think that all this is compatible with the lack of paradigm uses of the relevant logical words in inferences. (I’ve substituted “uses” for “applications” given that we’re dealing with inferences). One might suggest, for instance, that the introduction rule, the modus ponens schema, and the law of non-contradiction can be denied without giving raise to any suspicion that the person is incompetent with the meaning of “and”, “if”, or “not”. However, if we keep in mind the very weak characterization of paradigm cases supplied above, this seems false. It seems highly plausible to say that faced with someone who denies one of those principles, our default attitude is to think that the person must be conceptually confused. On the other hand, if I’m right in my previous remarks, this is not what we observe in the moral/normative case. So even assuming that something like *IC** holds and is common ground among the disputants, it is unclear how it could help the invariabilist.

plained in terms of a malfunction in his faculty of intuition, in the same way in which we explain, say, perceptual errors in terms of a malfunctioning in one of our five senses. Second, one could give up on the project of explaining the error: the error is *inexplicable* or *primitive*. What is a primitive error? Wright (2001) has a helpful answer: an error is primitive if it is not a “cognitive mistake”, where the latter is a malfunction “of a kind that might be appreciated as such independently of any imperfection in the result”; in other words, an error is cognitive if there is “some independently appreciable failure in the representational mechanisms” (p. 58).

The idea that we should appeal to “non-cognitive mistakes” seems to be implicitly assumed by some non-naturalists. If we opt for this solution, we could say that the outsider has a misleading *pretheoretical moral intuition*—to put things picturesquely, he’s under the influence of a “moral evil demon”: “something that causes moral error in a way that makes that error undetectable by ordinary means to the one who is deceived” (Wedgwood (2010), p. xx). Davis (2014) defends a similar view under the name of “divergentism”: “you can be epistemically faultless—that is, possess all relevant information and avoid all malfunctions in belief-formation—and still be mistaken about the truth of p ”. The idea is that the following situation is possible: A and B are in epistemically ideal conditions, A judges that p and B judges that $\neg p$; A is right and B is wrong. How can this be? Wedgwood (2010), as mentioned, talks of “evil demons”, Davis (2015) of “bare mistakes”, “weak falsehoods” and “failures of judgment”.

I think it is undeniable that both explanations require *very weighty* theoretical commitments. The most immediate problem is perhaps one of explanatory simplicity. Why should we appeal to defective intuitions rather than simply give a psychological characterization of our attitudes towards the normative outsider? It is certainly open to a variabilist to treat these subjects as *psychologically* or *sociologically* anomalous, where this doesn’t involve saying that they are semantically anomalous, that they aren’t “playing the moral game”, or that they are unlucky, so to speak, and simply lack the correct intuitions.

The foregoing considerations against the non-naturalist are far from being deci-

sive, and are unlikely to persuade someone who has independent reasons for being non-naturalist. However, to use Enoch's (2011) metaphor of "scoring points in an explanatory game", it's hard to deny that postulating *sui generis* mistakes in the envisaged way comes down to losing "plausibility points". My remarks can be strengthened by the usual metaphysical complaints against the "queerness" of irreducibly normative truth-makers. Even if one grants the possibility of "intuitional errors" (be they cognitive or non-cognitive), there must be something out there in the world in virtue of which the outsider is mistaken. The most powerful argument against the existence of this sort of entities is what Schroeder (2014) calls "the explanatory argument from supervenience". In broad strokes, it goes as follows. Everyone (or almost everyone) in the debate, non-naturalists included, accepts that *normative* properties supervene on *natural* properties: no two actions can differ in their normative properties unless they differ in their natural properties. Borrowing a famous example of Hare's, if someone said that a certain man X possesses exactly the same descriptive properties as St. Francis, but the latter was a good man and the former is not, we would be left wondering what he could possibly mean by his words (cf. Hare (1952) p. 145). But how can we make sense of this if the normative properties and natural properties are metaphysically distinct properties? Non-naturalism seems to be in violation of the so-called "Hume's dictum", according to which there aren't necessary connections between distinct existences. The non-naturalist has to posit *brute* metaphysical connections between normative and natural properties, and if this doesn't make the relation wholly mysterious (see, e.g., Horgan and Timmons 1992), it certainly constitutes a weighty explanatory burden for the non-naturalist (cf. McPherson, 2013).¹⁶

1.3.3 A Different Suggestion

If we don't ascribe linguistic incompetence based on the items to which someone applies or doesn't apply a certain term, how, if at all, do we count users of normative language

¹⁶There are slightly different versions of the supervenience argument against non-naturalism, according to which normative supervenience entails property-identity, and consequently there is no such thing as an irreducible normative property (see, e.g., Jackson (1998), Gibbard (2003b), Streumer (2008)). These arguments, though, unlike the one presented in the text, crucially depend on the metaphysically non-obvious thesis that necessarily coextensive properties are identical.

as incompetent? For reasons that will become clear soon, I shall focus on *moral* claims.

Consider the following:

(3) Eating fish is morally wrong, but I lack a desire not to eat fish/I'm not motivated not to eat fish.

Does an utterance of (3) constitute *prima facie* evidence of semantic confusion, in exactly the same way in which (4) does?

(4) Vin Diesel has zero hairs on his head, but he's not bald.

I'm inclined to answer in the negative. There seem to be many conversational contexts in which someone can use (3) to make clear that he's not motivated by considerations that one somehow acknowledges, shares, or regards as correct. The person may suffer from some mental condition, like addiction, apathy, depression, or some emotional disturbance. I think it is crucial to observe, though, that an amoralist is not the same figure as Hare's inverted commas guy, who makes purely descriptive claims concerning what is good or bad according to the standards that have currency in his group or society. In other words, amoralists and immoralists do bear some conative relation to the relevant action-type or standard. In light of this, a better candidate is perhaps the following:

(5) Eating fish is morally wrong, but I'm in no way against eating fish/I have nothing whatsoever against eating fish.

An utterance of (5) is much more likely to raise the suspicion that the speaker doesn't know what "morally wrong" means. One potential problem comes from the conceptual coherence of the question: "Why should I be moral?" After all, someone might recognize that morality requires so-and-so, but deny that there are reasons to act on those demands.¹⁷ Doesn't this show that (5) can be uttered in a totally felicitous way? I'm not sure, because I'm inclined to view the amoralist in question as someone who does have some con-attitude to eating fish, something which is overridden by further considerations. For this reason, (5) strikes me as incoherent. However, as the debate between internalists

¹⁷See Brink (1989) for an illuminating discussion of this issue.

and externalists show, intuitions are somewhat split between the contenders. I suggest we sidestep this problem by saying, in analogy to what we said above for descriptive predicates, that if someone utters (5) there is an initial temptation to infer that the speaker is conceptually confused. In other words, if someone says that x is morally wrong, but denies that by so doing he is presenting x in a negative light, or is committing himself to having some con-attitude towards x , then this constitutes some evidence of linguistic incompetence with “morally wrong”.

Things are probably more straightforward if we look at all-things-considered normative claims.

The upshot is this: if someone who is in apparently good psychological conditions denies that he has any con-attitudes (e.g. disapproval) towards a certain action that he claims to be wrong, our initial reaction is to think that he doesn’t know what “wrong” means. This is compatible with the subject being a competent user of the term, insofar as he has some reason for rejecting the relevant entailment.¹⁸

So, it is very important to point out that, if we stick to the case of morality, the claim defended in this section doesn’t require the truth of what we may call *conditional internalism about motives*, i.e. of the thesis that it is a conceptual truth that moral considerations provide some motivations to those agents who are in psychologically good conditions. Nor does it require the truth of *internalism about reasons*, i.e. of the thesis that it is a conceptual truth that moral considerations provide agents with some reason for action.¹⁹

Now, it is important to observe that if the remarks I’ve made concerning our ascriptions of competence/incompetence in the moral case are correct, then this is something that variabilists and invariabilists alike should predict. The difference, of course is that the invariabilist, unlike the variabilist, cannot be content with *just* that.

¹⁸A good feature of this account is that it gives a story that exactly parallels the way in which we ascribe incompetence, and retract previous ascriptions of incompetence, to users of descriptive expressions.

¹⁹Note also that both theses are formulated in their *weak* versions, insofar as what is required is *some* or a *pro tanto* motivation/reason, rather than a *conclusive* or *overriding* motivation/reason.

1.3.4 Conclusion

The central argument offered in this section builds on an asymmetry in our ascriptions of semantic competence/incompetence to users of descriptive vs normative language. I have used this asymmetry to support the claim that normative terms lack a positive and invariant extension. I considered some ways in which the invariabilist could resist the conclusion—namely, by denying the asymmetry or by explaining the asymmetry compatibly with their view. I've argued that neither option is particularly attractive. On the other hand, the variabilist can predict the relevant asymmetry, while not undertaking any of the problematic commitments that plague the invariabilist proposal.

1.4 Variability and Reducibility

Invariabilism, recall, is the view there is a uniquely correct and complete system of norms, or a small set thereof. As such, modulo some qualifications, it entails that if A assertorically utters a normative sentence φ and B assertorically utters its negation, it may be that one and only one of them is at fault, namely the one who doesn't accept the uniquely correct system of norms.

One component of my argument against invariabilism was the claim that full knowledge of the non-normative properties of a certain action x is compatible with competently applying or failing to apply any given normative predicate F to x . This is, in its essentials, premise **(P2)**. Since this claim has to do with *semantic* faults, we can call it *semantic variabilism*. As my previous considerations should have made clear, I think that semantic variabilism captures an important feature of normative terms vis-a-vis some other descriptive terms, and one that any satisfactory semantics should account for. It will prove useful to say a little more to clarify the content of the claim (i.e. the semantic variabilist thesis).

Moore (1903) famously claimed that normative terms cannot be *reductively analyzed* into non-normative ones. Let us understand this in the most liberal way possible, i.e. as

including both *definitions* and one-way analytic *entailments*.²⁰ The claim is that there are no analytic entailments from statements that don't contain normative vocabulary to statements that do contain such vocabulary. Call this thesis *semantic non-reductionism*.

While this may be obvious, it is important to stress that semantic variabilism and semantic non-reductionism are two distinct claims. More precisely, while the latter arguably entails the former, the converse doesn't hold. Let me explain why the converse entailment doesn't hold by means of two examples. The first is Lewis' analysis of value in his (1989). Lewis analyses " x is a value" in terms of " x is such that we would desire to desire it under conditions of full imaginative acquaintance". However, the *definiens* contains the term "we", which as Lewis himself observes, allows for a range of more or less inclusive readings—"we" could refer to all mankind, or to "just me", i.e. to the speaker only (and, of course, there are many other versions in between). The relativistic version of Lewis' reductive analysis is a form of semantic variabilism.

The second example is Finlay's reductive analysis of "ought" in terms of ends and probability (see Finlay (2009) and (2014)). Roughly, the idea is that "A ought to ϕ " really means: "(Some implicit) end E is more likely to obtain if A ϕ s than if A does anything else instead". Again, different people can have different ends, so this is obviously compatible with semantic variabilism.

1.5 Natural Language Semantics and Metanormative Inquiry

Broadly speaking, the job of a semantics for a natural language L is to characterize the conventional meanings of the expressions of L .²¹ One natural way of thinking about

²⁰The notion of *analytic entailment* is meant to be broader than (and thus to encompass) that of *analytic definition* or *synonymy*. $G(x)$ may be an analytic entailment of $F(x)$ regardless of whether being F can or cannot be defined in terms of being G , or of whether ' F ' is or is not synonymous with ' G '. This is of particular importance for those who adopt a notion of synonymy which is so fine-grained that 'no two different words have the same meaning' (see Goodman (1949), p. 6) or fine-grained enough that 'male sibling' and 'brother' don't count as synonyms (cf. Carnap's notion of 'intensional isomorphism', in his (1947).

²¹I'm using "conventional" in contrast to "conversational", in the Gricean sense. So, "conventional meaning" refers to those aspects of "meaning" that are encoded in the lexicon or grammar, instead of being calculable on the basis of the Gricean cooperative principle and its maxims.

the object of semantics, i.e. conventional meanings, is in terms of an abstract model or representation of what, under certain ideal conditions, a competent speaker of the language knows, understands, or grasps.²² The view that semantics is, essentially, a formal explication of linguistic competence is, arguably, the standard one.²³ Here is one recent statement of it:

[...] semantics seeks to provide an account of what a speaker understands when they grasp the meanings of their sentences, i.e. what they know when they know what their sentences mean (Glanzberg (2015)).

One important corollary of this approach is that speakers' judgments or intuitions about the properties of, and relations between, linguistic expressions, under certain ideal conditions, are reliable indicators of the conventional meanings of these expressions. In other words, the idea is this: the right semantic value to assign to a certain expression is the one that generates the correct predictions concerning language use, namely the one that best captures speakers' judgments/intuitions (again under certain ideal conditions), about the acceptability of certain constructions, the entailment relations among various sentences, and the truth/correctness of certain utterances (cf., for instance, Yalcin (2014) and Chierchia (2000)).

Something like this is implicit in a good deal of old and new metanormative theorizing. For instance, Moore (1903) famously claimed that sentences like: "*x* is good, but is not pleasant" do not sound self-contradictory, and that this shows that "good" and "pleasant" are not synonymous or don't have the same linguistic meaning.²⁴ Error theorists and their rivals debate on whether there is linguistic evidence for the thesis that "the basic, conventional, meaning of moral terms" incorporates "a claim to objectivity"

²²I'm taking semantic properties to be those associated with an expression-type by linguistic convention. On this understanding, a conventional implicature counts as semantic. Of course, if one understands semantic properties as those that contribute to truth-conditions or "at-issue content", then "semantic" and "conventional" turn out to have disjoint extensions, and a conventional implicature will count as a pragmatic feature.

²³Though it is not universally shared. On a different view, even full competence only puts one in a position to know the proposition asserted/communicated, and doesn't discriminate between what is conventionally encoded in the sentence and what is not. For a statement of a view like this, see Soames (1989) and (2008).

²⁴Actually, Moore claimed much more than that. But that's not relevant for now.

(Mackie (1977), p. 35). Relativists argue that contextualists can't explain why someone can felicitously use expressions of linguistic denial ("no", "nuh-uh", etc.) in response to moral assertions that are incompatible with the moral standards that he or she accepts (see, for instance, Kölbel (2004), and MacFarlane (2007)).

The idea is that based on considerations like these we should be able to give a semantics for normative terms, i.e. an account of their conventional meaning. According to the dominant view in semantics, namely the *truth-conditional* paradigm in its *model-theoretic* version, the semantic content of a declarative sentence (in context) is a truth-condition determining proposition and the semantic content of the subsentential expressions consists in their contribution to the overall sentential meaning.²⁵ On this conception, semantics employs propositions and a proposition is, or at least determines, a set of possible worlds. Hence, the semantic content of a sentence (in context) determines the set of possible worlds in which the sentence-context pair is true. (Further issues about the metaphysics of propositions—for instance, about whether propositions are just *sets of worlds*, or more *structured entities*, or something else altogether—can be bracketed for now).

Within this paradigm, one especially influential approach is the so-called "double-index" framework, as developed by Kaplan (1977), Lewis (1980), and others. Two types of meaning, in the intuitive sense of "meaning", play a role in Kaplan's semantics:²⁶ a *character* and a *content*. A character is a function from a context to a content. In other words, it is a rule that tells you what the content of a certain expression is in any possible context in which it may be used. It is the conventional or linguistic meaning of a certain expression, that you know when you know that, say, "I" means something different from "now", and what is stable across all possible uses of "I". A context typically includes a world, an agent, a location, and a time. A content is (or at least determines) a function from a circumstance of evaluation/index to an extension. At the level of sentences, this function is also called an intension or a proposition and is taken to be what a speaker conventionally says (communicates, asserts, claims, etc.) when using the sentence in a

²⁵For now, I will abstract away from other conventional, and yet non-truth-conditional features, of expressions.

²⁶For the most part I shall use Kaplan's framework and terminology.

context. A circumstance of evaluation typically includes a world (or a world-time pair, as in Kaplan).²⁷

Furthermore, it is generally acknowledged that many natural language expressions have variable (or non-constant) characters and/or contents, i.e. characters that deliver different contents in different contexts and/or contents that determine different extensions in different circumstances. Many, but not all. For example, the character of “taller than 1.85” is plausibly constant; likewise, the content of “prime number” is a constant function which takes every world to the same set.

To recap, according to a fairly standard picture a semantic theory assigns a single semantic content to each sentence-in-context. This content is a proposition and corresponds to what a speaker says/asserts when using the sentence in context. The same proposition may be true at a world w and false at a world w^* . Now, consider an utterance of the following sentence by Jim:

(5) Mia saw a bat yesterday.

Of course, “bat” has (at least) two possible meanings—it can refer to a flying mammal or to a stick used to hit. I will return to this point below. Let us assume that the correct disambiguation is “flying mammal” and thus let’s reformulate (5) as follows:

(5*) Mia saw a bat_{*m*} yesterday.²⁸

When is an utterance of (5*) true or correct? Plausibly, language users will consider utterances of (5*) as true if the content of (5*) at the context associated with the utterance is true at the world of that context (other possible worlds, i.e. other ways the world might have been, are irrelevant to our intuitions about the truth/correctness of Jim’s utterance).

It is for this reason that Kaplan provides the following principle for connecting the technical, formal-semantic, notion of truth at a context and index to the ordinary, intuitive, notion of truth *simpliciter* (at a context):

²⁷One can depart from the orthodoxy by proliferating the coordinates in the index, but also by doing away with the index. Lewis’ *octuple* of parameters in “General Semantics” (1970) is a clear instance of the former strategy, while Schaffer’s (2012) *index-free* semantics exemplifies the latter.

²⁸I’m using the subscript to disambiguate. I’ve rewritten (5) as (5*) just for illustrative purposes. I’m not assuming, for instance, that there are two distinct words/sentences that are phonetically and orthographically indistinguishable.

(**K**): A sentence-context pair $\langle s, c \rangle$ [an “occurrence” of s in c , in Kaplan’s vocabulary] is true iff the content expressed by s in c is true at the circumstance of evaluations determined by c .

Principles like (**K**) are the *postsemantics*, rather than the semantics proper, in MacFarlane’s vocabulary. In his (2014), MacFarlane helpfully articulates three components of a theory of meaning: the *semantics*, the *postsemantics* and the *pragmatics*. The job of the first is to give a definition of truth at a context and index; the job of the second is to give a definition of truth-at-a-context (or at-contexts, if one uses contexts of assessments in addition to contexts of use) in terms of the definition of truth at a context and index; finally, the pragmatics “relates the output of the semantic theory to the use of sentences” (p. 55), by specifying when it is correct to use these sentences. Since it is the notion of truth-at-a-context (or at contexts), rather than “the technical notion of truth at a context and an artificial sequence of coordinates” (p. 57), that is pragmatically relevant, the postsemantics—precisely by providing a definition of truth for sentences (at a context of use, or at a context of use and context of assessment) in terms of the notion of truth used by the semantics proper—bridges the gap between the semantics and the pragmatics.

To complete this picture, it is important to keep in mind the distinction between a *concrete utterance* on the one hand, and the *abstract sentence-context pair* which is taken to correctly represent it, on the other. It is not the job of a formal semantic theory to decide which out of a range of candidate abstract sentence-context pairs correctly represents a concrete utterance. More precisely, the semantics assumes as its inputs: (i) a mapping from a string of (types of) sounds or of marks to a linguistic meaning and (ii) a mapping from a concrete speech-situation to an abstract representation thereof, which fixes the values of the contextual parameters. No formal framework will tell you whether, say, $\langle s_1, c_1 \rangle$ or $\langle s_2, c_2 \rangle$, where s_i is a certain base language sentence (i.e. a sentence disambiguated as in (5*)), and c_i is a certain abstract context, correctly represents a concrete utterance. How to deal with these sorts of ambiguities or indeterminacies is an open question. One could think that they are better represented at a *presemantic* stage. Consider, for instance, lexical ambiguity. One possible view is that an expression

is ambiguous iff it can be characterized by two distinct *logical forms*.²⁹ Alternatively, one can instead think that the semantics itself should be enriched, perhaps by assigning sets of meanings to ambiguous expressions (cf. Poesio (1996)).

To conclude this section, it will prove useful to rehearse a distinction between three broadly meta-normative inquiries, which we can call the *semantics*, the *metaphysics*, and the *metasemantics* of normativity.

The first, i.e. semantics, addresses the question: what is the semantic value—in the sense of “the character” and “the content”—of a certain normative expression ε ? Some of the views I have mentioned at the beginning—contextualism, relativism and expressivism, for instance—are typically construed as views about the *semantics* of certain expression.³⁰ More generally, the framework introduced in this section is mostly a framework for theorizing about the semantic properties of natural language expressions.

There is next a metaphysical question: what *worldly property/properties*, if any, does ε stand for? Or yet, what are the *truth-makers* of normative judgments/assertions? Of course, if one thinks that the semantics involves Russellian propositions—i.e. structured complexes of objects, properties and relations—the relationship between the semantic and the metaphysical enterprise will be much more tight. Regardless, though, one will face the choice between different worldly properties that “wrong” might denote: e.g. a mind-dependent property, a natural (mind-independent) reducible or irreducible property, a non-natural property, or no property at all.

(In the literature, people often speak of “properties” both in the sense of semantic values, or linguistic meanings, of predicates and in the sense of truth-makers or real features of object. While one might have reasons for thinking that the one and the same entity should play both roles, it is important not to conflate the two.)³¹

²⁹ Authors in the generative grammar tradition typically follow this route. More precisely, they assume that there is a *base language* or *logical form language*, a *surface language* and a set of *transformational rules* which transform the base language sentences into the surface sentences. See, e.g., Lewis (1970) and Gillon (1990).

³⁰“Semantics” here can be understood in a broad sense, including the postsemantics and the pragmatics, as defined above.

³¹As Tooley (1989, pp. vii-viii) rightly notes, there is “no one-to-one mapping between properties in the sense of entities that ground objective sameness and difference in the world, and properties in the sense of entities that serve as semantic values of predicates and the use of the term ‘property’ in these two very different senses is unfortunate, since it tends to obscure the crucial difference”. In this connection,

Finally, there is a *metasemantic* investigation, which is instead concerned with the question: in virtue of what does ε have the semantic value that it has? A number of options are available: Descriptivists say it is a folk-theory about ε -things, Causal-historical views say it is some casual connection, Gricean mentalists say it is language users' mental states, etc.

following Lewis (1986), people sometimes distinguish between an *abundant* and a *sparse* conception of properties: "The abundant properties provide the semantic values of meaningful predicates, while the sparse properties carve out the joints of nature on which the causal powers hinge" (Schaffer 2004, p. 92).

Chapter 2

What Normative Words Mean: Some Old and New Stories

2.1 Introduction

Variabilism is an attractive view: it nicely explains a certain asymmetry in our attributions of semantic incompetence to users of normative as opposed to descriptive terms, and it does so without introducing a problematic epistemology and ontology. It is thus interesting to see how several competing semantic accounts can accommodate it. Sections 2.2 to 2.6 outline four such accounts—*Contextualism*, *Expressivism*, *Relativism*, and the *Error Theory*¹—and raises some problems for each of these views. This sets the stage for the introduction of my own positive proposal, in Chapter 3.

¹Strictly speaking it's inappropriate to describe the error theory as a distinct *semantic* option. The error theory simply says that all substantive normative statements are untrue, and different theories of the semantic contents of normative sentences (including contextualism, relativism and expressivism) can in principle make this prediction. On the other hand, it is certainly true that the error theory is typically cast in terms of an absolutist semantics, i.e. a semantics according to which a normative sentence expresses the same possible world proposition across contexts, and the truth-value (or lack thereof) of this proposition is determined by the world.

2.2 Contextualism

2.2.1 Introducing Contextualism

On a contextualist semantics, normative sentences are implicitly relative to a contextually supplied parameter, like a standard, framework, ordering, end, body of norms, etc.² The content of normative expressions, thus, depends on features of the context in which the sentence is uttered.³

To a first approximation, the idea is that a normative sentence like “eating fish is wrong” means something like: *eating fish is wrong according to standard so-and-so*, where the context of utterance supplies the relevant standard.⁴ This is compatible with different views about *which standards* matter. It may be the standard of the speaker, on the model of “my standard”; or it may be the standard shared by the conversational group, on the model of “our standard”; or yet, the standard parameter may be determined by the conversational context in a more flexible way, on the model of “nearby” , “local” or “left”—“eating fish is wrong” would mean something like: *eating fish is wrong according to the contextually relevant standard* (which may be the one of the speaker’s, the one of a group, or an external standard, i.e. a standard not necessarily endorsed in the conversational context). The last route is, arguably, the most popular among contextualists,⁵ although it is safe to say that, for most of them, the *central uses* are those relativized to the speaker’s standard.

At a fairly abstract level and simplifying a bit, I will take a standard, e.g. a moral

²Different authors suggest different relativizations. I will assume a relativization to standards. However, for my current purposes, nothing should turn on this choice. So, you can think of “standard” as a place-holder for whatever it is that you think the content of normative sentences is relative to. For example, if you are a particularist (cf. Dancy (2004)), and you think that there are no general normative standards (or, more weakly, that general principles do not play any fundamental role in morality) you can introduce situation-specific standards or perhaps relativize to something that has more sensitivity to context built into it—e.g. interests or ends (as in Finlay’s (2014) end-relational semantics).

³Contextualist theories of normative language are defended, among others, by Harman (1975), Wong (1984), Dreier (1990), Copp (1995), Phillips (1997), Finlay (2009) and (2014), Dowell (2013).

⁴There are at least two different ways of understanding the context-sensitivity of normative terms: one can view them as *indexicals*, in line with Kaplan’s semantics for “I”—their semantic content depends on the context; alternatively, one can take them to be *incomplete* or *relational* predicates, in line, e.g., with Kratzer’s (1977) and (1981) semantics for modals—the semantic content is invariant, but the term has an additional argument place (which, in turn, may or may not be syntactically realized as a free variable). These issues are largely orthogonal to my goals here, so I will ignore them hereafter.

⁵See, for example, Dreier (1990), Dowell (2013) and Finlay (2014).

standard, to be (or to be something that determines) a complete and consistent assignment of truth-values to moral sentences (cf. Chapter 1, §1.1). I'm individuating standards simply in virtue of the truth-value assignment that they determine. Since *subjects* generally have standards, we can also pick them out by means of descriptions like “the speaker’s standard”, “the audience’s standard”, “the standard shared by the speaker and the audience”, and so on. This doesn’t mean, though, that normative claims are *about* the subjects’ (current or idealized) conative attitudes, and thus *describe* or *report* these attitudes. Most contemporary contextualists, even when they are *speaker*-contextualists, are not *subjectivists*. That is, they do not think that “eating fish is wrong” has the same character as “I disapprove of eating fish” or “eating fish is disapproved of by me”. A subject’s desires, beliefs and dispositions play a role in determining which standard that subject has; and a given subject’s standard may in turn be what is contextually identified, thereby determining what proposition is expressed by the sentence. However, this proposition is not about the subject’s conative states, at least not in the way suggested by the subjectivist.⁶

In Chapter 1, I characterized contextualism, relativism and expressivism as variabilist accounts. This is how these views are generally presented in the meta-ethical literature. However, it is important to keep in mind that all these frameworks can be understood, to adapt a slogan from Harman (1975), as soberly *linguistic* or *semantic* theses.⁷ If so construed, they are not committed to being variabilist. Let me explain. There are, I think, three claims commonly associated with meta-normative contextualism:

(C1) Normative sentences have inconstant characters and are implicitly relative to a parameter, i.e. a standard, supplied by the context of utterance;

(C2) The same normative sentence may be interpreted with respect to different

⁶On this point, see (Dreier’s (1990), pp. 14-17) objection to Michael Smith’s version of subjectivism. Most contemporary contextualists are not subjectivists. See, for instance, Finlay (2014) and Evers (2011). That said, there remains a broad sense in which according to contextualists normative claims do report attitudes, a sense that corresponds to the one in which “‘The present King of France is bald’ reports the existence of the King as well as his baldness” (Jackson and Pettit (1998), p. 239) even though the existence of the King is not the focus of attention.

⁷(Harman (1975), p. 3) famously presented his relativism (in our terminology, his contextualism) as a “soberly logical thesis—a thesis about logical form, if you like.”

standards when used by different speakers;

(C3) There isn't a uniquely correct or valid normative standard.

These three theses are conceptually independent, and worth separating. To adopt a contextualist semantics for normative terms is to subscribe to (C1). But one can adopt (C1) and be a wholly traditional invariabilist by denying either (C2) and (C3), or just (C3).

In the former case, one will say that the context of utterance always supplies the same universally correct standard—or, at least, that it sets some constraints on what standards are acceptable.⁸ In the vocabulary introduced in Chapter 1, invariabilism would be characterized not at the level of semantics, but of *metasemantics*, in particular of what King (2014) calls the “the metasemantics of context-sensitivity”.

In the latter case, i.e. if one denies just (C3), one will say that the same utterance may invoke different standards, and thus that two utterances of syntactically contradictory sentences may both be true. One will add, though, that the two claims will not be equally valid or correct: one standard is objectively better than the other; if you like, there is something like the context one ought to be in. Invariabilism would thus be formulated as a thesis at the level of what is *pragmatically* conveyed, in a way that is consistent with holding on to (C1).

For brevity, I shall stipulatively use “contextualism” as a shorthand for variabilist contextualism, i.e. for the conjunction of the three claims. (Whenever I need to refer to contextualism as a thesis about the semantic content, or compositional semantic value, of normative sentences, i.e. as thesis (C1), I will make that explicit, if there is any risk of misunderstanding).

2.2.2 Virtues and Vices

I will focus my attention on what is arguably the most elaborated defense of contextualism to date, namely Finlay's view, as developed in his (2009) and (2014).

⁸As an example, consider Copp (1995) who relativizes the content of moral sentences to “relevantly *justified* or *authoritative* moral standards”; or, more recently, (Ridge's (2014) p. 40) relativization of practically normative claims to “*acceptable* standards of practical reasoning”.

An advantage that can be claimed for Finlay’s specific variety of contextualism is that it provides a *uniform semantics* for “ought” and “good” in all their uses (including non-moral ones). It avoids having to say that “ought” and “good” are systematically ambiguous. This is important, if we consider that words like “ought” and “good” may receive a range of interpretations, depending on the context in which they occur.

This is especially clear in the case of modals (like “ought”, “must”, “may”, etc.). Consider “ought”. The local weather forecaster is not making a moral statement when she says *that a rainbow ought to be visible in the afternoon*. In fact, she is not making a normative claim at all, i.e. she’s not talking about any norms or values. Intuitively, he’s making a claim about what is likely to happen in the light of the evidence (at least, in the light of the best evidence she possesses).

You might think that the “ought” that figures in the meteorological claim is not the same “ought” that figures in a moral claim like *that you ought to keep your promise*. On this line, “ought” would be ambiguous between agential readings (like the latter, moral one) and non-agential readings (like the former, epistemic one). This is a view that many philosophers sympathize with. But let us set aside predictive or epistemic uses of “ought”, and let us fix on its normative uses. “Ought” can figure in instrumental or “means-to-ends” claims, as well as in statements about what one morally/aesthetically/legally/prudentially/epistemically/all-things-considered, etc. ought to do, feel, believe, etc. It is plausible to suppose that in all these normative uses the term retains a *common semantic core*, rather than being systematically ambiguous, in the way in which “bank”, “bat” or “couch” are. Thus, even if “ought” was ambiguous between normative and non-normative readings, it seems hopeless to posit a bunch of distinct lexical entries for each of the possible normative readings of “ought”.⁹

This conjecture motivates the canonical Kratzerian semantics for modals.¹⁰ The main idea is that modals are context-sensitive (existential or universal) quantifiers over possibilities. More precisely, modals are interpreted with respect to two “conversational

⁹The view that “ought” is ambiguous between epistemic and normative uses is defended by Geach (1982) and Schroeder (2011). Finlay (2009) and (2014) opts instead for a unified semantics of all normative and non-normative uses of “ought”.

¹⁰See Kratzer (1977) and (1981).

backgrounds”, i.e. two parameters supplied by the context: a *modal base* f and an *ordering source* g . The modal base determines the range of possible worlds to be considered, and the ordering source induces a ranking of the worlds in the modal base.¹¹ The two parameters determine the flavor of the modal (metaphysical, epistemic, instrumental, deontic, etc.).¹²

To sum up, by assuming that modals always denote a relation between a proposition and some background assumptions, contextualism can offer a unified semantics for modals, i.e. one that doesn’t posit lexical ambiguities between normative and non-normative readings of “ought”, nor between its instrumental and categorical uses, nor yet between its various categorical uses (e.g. practical and non-practical).

While I’ve focused on auxiliary modals, similar considerations hold for terms of other syntactic categories, for example adjectives like “good” and “right”.¹³ There seems to be little hope for a theory according to which “good” shifts its meaning when it occurs in non-moral contexts, i.e. in sentences like: “this is a good strawberry”, as opposed to moral contexts, i.e. in sentences like: “cutting taxes is good”.

To summarize, a plausible semantics for our trio of terms (“good”, “ought”, “right”) should explain why their conventional meaning is constant across all their normative (and perhaps also non-normative) uses. The brand of contextualism defended by Finlay offers a neat explanation.¹⁴ This may constitute an advantage over alternative semantic proposals. For example, on the plausible assumption that instrumental uses of “ought” need not express the speaker’s conative attitude, expressivism would need a mixed semantics, i.e. a semantics which is partly expressivist and partly contextualist, to accommodate those uses. Moreover, it is unclear whether expressivism can preserve the uniformity of the character of “ought” across different normative uses (moral, legal, aesthetic, etc.). Plausibly, the expressivist should identify a common core in the possibly different con-

¹¹More precisely, f and g are functions from worlds to sets of propositions, call the latter F and G , where F represents the relevant circumstances and G represents the demands of the relevant ordering/standard.

¹²In his end-relational semantics for modals, Finlay articulates a simpler semantics that doesn’t posit an ordering source parameter (see Finlay 2009 and Finlay (forthcoming)).

¹³For a thorough discussion of “good”, see Finlay (2014), pp. 19-47.

¹⁴The same can be said of Phillips’ (1997) contextualist proposal.

tive attitudes expressed by the *oughts* of morality, legality, rationality, etiquette, and so on. This is no trivial task. An expressivist might hope to avoid both problems by going metasemantic. According to *metasemantic expressivism*, as developed by Ridge (2014), the meaning of “ought”-claims is captured by an orthodox truth-conditional semantics; however, in *some* of its uses, “ought” has its actual meaning *in virtue of* expressing some non-cognitive mental state; in some other of its uses, “ought” has its actual meaning in virtue of expressing some cognitive, representational, mental state. However, metasemantic expressivism trades a mixed semantics for a mixed metasemantics. This is a theoretical cost: arguably, *meaning-fixing facts* should not be multiplied beyond necessity anymore than *meanings*.

A second important point is that contextualism can allow for at least three possible uses of moral terms which may be more problematic on other accounts, and which I shall call *external*, *cold*, and *reverse* uses respectively. External moral statements, which correspond to Hare’s (1952) famous “inverted commas” claims, are those in which the speaker doesn’t, in any way, portray herself as affectively/emotionally related to the standard required for the truth of her claim. There seem to be conversational contexts in which we can say that *A ought to do such-and-such*, given the moral standard she accepts, without endorsing in any way that standard ourselves. (Think about two friends discussing whether a certain emperor, king or president ought or ought not to have done *X*; or think about a non-vegetarian saying to a vegetarian that he should not eat a certain meat dish). Further, a moral error theorist, i.e. someone who denies that anything is wrong, right, etc., may be able to give “moral” advice to a friend (by assuming the perspective of the relevant individual) or to work out the implications of the moral standards operative in his society; and a parent may well opt for teaching his child the “conventional morality”, even if (deep down) he finds it largely flawed and doesn’t approve of it in the relevant way. These uses are probably rare, but not impossible. One could go for a “mixed view” on which the relevant terms have a different meaning than the usual one when used externally. As always, though, positing ambiguities should be a last resort.

Now, recall the contextualist schema given above, according to which “x is wrong” means, roughly: *x is wrong according to the standard of the context*. This allows for the possibility that a certain standard, which the speaker doesn’t necessarily subscribe to, is made salient or relevant in the conversation. Importantly, it is open to a contextualist to say that external uses are infelicitous, or in any case are to be treated differently from internal uses. This sort of contextualist could then say that the argument of the character-function associated with “wrong” is simply the speaker’s actual motivational states, rather than a more complex function which can take different arguments.

Cold and *reverse* statements, on the other hand, are statements that overtly disavow desire-like or advice-like features (“Eating fish is wrong, but I don’t disapprove of it/ I’m not saying that you should not eat fish”) or that are accompanied by reverse desire-like or advice-like features (“Eating fish is wrong, and/but I approve of it/I recommend that we go fishing together”). Unlike the former, cold and reverse normative assertions are *internal* rather than *external*.¹⁵ They are used to communicate that the speaker is not motivated by, or is motivated by in contrary ways, by standards that he himself somehow acknowledges, shares, or regards as correct. While these uses are somewhat puzzling, I take it that we hesitate in ascribing semantic incompetence to subjects, like the amoralist and the immoralist, who make cold and reverse assertions respectively—at least, when given the right contextual clues.¹⁶ If this is true, these uses create troubles for any analysis on which an (internal) utterance of a moral sentence *conventionally* communicates that the speaker has a conative attitude *of a certain kind* (i.e., motivational), or is conventionally prescriptive in a certain way. This doesn’t mean that we should dismiss out of hand the view that, in their internal uses, moral assertions express *some kind* of conative attitude by virtue of their conventional meaning. However, as I shall explain below, it makes trouble for some traditional expressivist models. Of course, one could concoct versions of expressivism that accommodate cold and reverse uses, but these strategies are likely to

¹⁵I borrow the internal/external vocabulary from Hart, who distinguishes between internal and external legal statements. The former are normative statements made from the point of view of an adherent of the legal system; the latter are statements of fact which an observer of the system might make even if he did not accept it. (cf. Hart 1961/94, p. 108).

¹⁶See the discussion in Chapter 1...

comport revisions that are in tension with the original expressivist program (again, I will return to this point below). On the other hand, contextualists, like other descriptivists, can more consistently say that the attitude is only *conversationally* communicated (cf. Finlay (2005), Strandberg (2012)) or, alternatively, can posit *weaker* conative attitudes (e.g. attitudes that don't guarantee motivation).

As before with external uses, it is important to note that only *some versions* of contextualism may be able to accommodate cold and reverse uses. For example, these uses are arguably not predicted by a view on which the contextually determined standard is fixed, say, by a certain attitudinal profile of the speaker.

Finally, contextualism is theoretically more conservative than its competitors. Unlike *expressivism*, it doesn't face some familiar difficulties for a *semantics of attitudes*, i.e. a semantics according to which the meaning of a normative sentence consists in the conative attitude it is conventionally used to express; unlike *relativism*, it doesn't encounter the problems that afflict views that countenance propositions with relative truth-values as the objects of mental states and speech-acts. (Of course, we shall see more about these problems in what follows).

Unfortunately, contextualism faces a number of well-known challenges. Specifically, it is contended that a contextualist semantics (i) has troubles representing the contents of attitude ascriptions/speech-reports,¹⁷ (ii) entails that apparently inconsistent claims are perfectly compatible, (iii) and can't capture the phenomenon of "retraction" (cf., among others, Kölbel (2004) and (2009) and MacFarlane (2014)). I will mostly focus on the second worry, the so-called *lost disagreement* problem, which is by and large the most discussed in the literature.

There is a set of alleged data associated with the contention that contextualists can't explain disagreement. And there is some discussion concerning which of the purported data is genuine. I will work with a very weak characterization of the *explanandum*, for I

¹⁷This objection is typically pressed in light of the Kaplanian assumption that natural languages (or, better, English) don't contain "monsters", i.e. operators that shift parameters of the context of use (see Kaplan (1977)). It is open to the contextualist to allow for monsters, i.e. to say that speech verbs and attitude verbs can shift the context (see Schlenker (2003)). However, this is somewhat in tension with the contextualist conservatist aspirations. See Kölbel (2015) for discussion.

think this is sufficient to make trouble for contextualism. It goes as follows:

Disagreement: *Typically* (i.e. in *most* cases), if two agents A and B sincerely and assertorically utter two syntactically contradictory moral sentences φ and $\neg\varphi$, they thereby disagree in the sense that A/B can felicitously reply to B/A by saying: “no”, “nu-uh”, “I disagree”, “that’s wrong”, “that’s false”, etc.

On this minimal understanding of “disagreement”, what calls for explanation is the felicity of a range of “disagreement markers” in response to a given normative utterance. Three points follow that are especially important to keep in mind. First, the characterization given doesn’t imply that the disagreement has to be “rationally sustainable” (see, on the contrary, Wright (2006)), or worth having, or any such thing. Second, **Disagreement** doesn’t discriminate between the two varieties of disagreement that MacFarlane (2007) and (2014) calls “non-cotenability” and “preclusion of joint accuracy”.¹⁸ In light of this, I think it is weak and coarse-grained enough to serve as a neutral explanatory desideratum. Third, **Disagreement** is not a universal statement, and is thus compatible with exceptions. Crucially, though, I take it that everyone has to explain why, generally or typically, B can express his disagreement with A’s utterance by using any of those constructions (or others that entail them).

Disagreement is easily explained by a theory, like relativism, that assigns the same semantic content, p , to utterances of φ , and mutually contradictory semantic contents, p and $\neg p$, to utterances of φ and $\neg\varphi$. Assuming that A believes that p and B disbelieves that p , such theory can predict the propriety of the various ways in which A/B can object to B/A.

By contrast, **Disagreement** is not easily predicted by a contextualist semantics. On most contextualist proposals, moral statements are, in general, implicitly relative to the speaker’s standard. This means that if Mia utters the sentence “eating fish is wrong” (1) and Jim utters its negation “eating fish is not wrong” (2), the semantic contents of (1)

¹⁸MacFarlane (2014) defines the two notions as follows. *Non-cotenability*: “I disagree with someone’s attitude if I could not coherently adopt that same attitude (an attitude with the same content and force) without changing my mind—that is, without dropping some of my current attitudes” (p. 121). *Preclusion of joint accuracy*: “The accuracy of my attitudes (as assessed from any context) precludes the accuracy of your attitude or speech act (as assessed from that same context)” (p. 129). The two notions come apart if, say, A and B who inhabit two distinct possible worlds, affirm and deny the same world-neutral proposition. Their attitudes are non-cotenable, but nothing prevents them from being equally accurate.

and (2) are perfectly consistent—the former expresses the proposition *that eating fish is wrong according to Mia’s standard*, while the latter expresses the proposition *that eating fish is not wrong according to Jim’s standard*. If so, why should they object to each other’s statements?

Contextualists have offered several replies. Almost all of them are of a *pragmatic* sort: they account for **Disagreement** without positing mutually contradictory semantic (and asserted) contents. I will briefly look at two approaches that exemplify the preferred pragmatic strategies pursued by contextualists: Björnsson and Finlay’s (2010) “quasi-expressivism”¹⁹ and Plunkett and Sundell’s (2013) “metalinguistic negotiation” account.

I will present some worries for each of them. My discussion will be fairly quick; I will return to the topic in Chapter 3, after having outlined my positive proposal.

Let’s work with our simple example:

- (1) Eating fish is wrong
- (2) Eating fish is not wrong

The quasi-expressivist strategy invokes disagreements in conative attitudes, like desires and preferences. The semantic contents of (1) and (2) are perfectly consistent—the former expresses the proposition *that eating fish is wrong according to Mia’s standard*, while the latter expresses the proposition *that eating fish is not wrong according to Jim’s standard*. However, utterances of (1) and (2) pragmatically convey inconsistent conative attitudes, like desires, preferences, etc. The solution is “expressivist” because it appeals to desire-like states (rather than to belief-like ones); but it is only “quasi”-expressivist because the attitudes don’t figure into the semantics, but are rather part of the pragmatics. More specifically, each of the speakers *conversationally implicates* that she has a certain attitude towards a certain action-type or event-type.

I think there is one main problem with this proposal. Disagreements in attitudes don’t always license the use of expressions like “no”, “nuh-uh”, etc. It seems utterly unnatural to use these expressions in reply to an utterance such as “I like Venice”.²⁰ In

¹⁹See also Finlay (2014).

²⁰This objection is raised, for instance, by McKenna (2014) and Silk (forthcoming).

response, advocates of the quasi-expressivist account have offered two considerations. The first, developed in Björnsson and Finlay (2010), has it that the use of the relevant disagreement markers is licensed by the fact that the “at-issue proposition” is expressible in the context by the same sentence. More precisely, “an evaluator might be able to respond felicitously by intentionally addressing a proposition related but distinct from the one the speaker asserted, by virtue of its greater relevance in the context” (Finlay 2014, p. 240). Finlay’s more recent (and preferred) solution²¹ is an appeal to *semantic blindness*: speakers fail to recognize the truth of contextualism, and thus mistakenly assume that there is a common propositional content (cf. Finlay 2014, pp. 241-243).

It may be that, in the end, contextualism (perhaps in its quasi-expressivist clothing) is vastly superior to all its competitors in every other way, so much so that we should simply accept a certain dose of semantic blindness. However, I think we should first consider whether there are equally good alternatives that make do without ascribing this sort of error to speakers.

A further problem is that the incompatibility between (1) and (2) seems to be grounded in the conventional meaning of the two sentences. As Kölbel (2007, p. 286) points out, if a person utters (1) and another utters (2) “we know without any further knowledge about their background beliefs that one can’t accept what the other asserts”. Moreover, the incompatibility arguably disappears if the indexical element is made explicit, as in “eating fish is wrong relative to the vegan moral code” and “eating fish is not wrong relative to the omnivore moral code”. In light of this, one may wonder whether a merely pragmatic-conversational explanation of **Disagreement** is sufficient to correctly capture the phenomenon.

The second pragmatic strategy is the metalinguistic/metacontextual approach developed, for instance, by Sundell (2011) and Plunkett and Sundell (2013). The idea is that even if (1) and (2) are logically compatible and may both be true, by making their assertions Mia and Jim pragmatically advocate for two competing parameter settings for

²¹Finlay now thinks that his previous solution “may be too cavalier about the role of the demonstrative ‘that’, or talk about what the speaker ‘said’ or ‘believed’” (2014, pp. 240-241). I share these worries, and I will thus only comment on the newer approach.

the context-sensitive term “wrong”—more informally, they disagree over what concept of *wrongness* is the best one to have, or over what context we ought to be in. This is, roughly, what a “metalinguistic negotiation”, in Plunkett and Sundell’s jargon, consists in.

By going metalinguistic rather than quasi-expressivist, one may be in a better position to accommodate the full range of disagreement markers. Consider a classical example of a metalinguistic dispute. Mia and Jim, who know the respective heights of Andrew and Mark, utter (3) and (4) respectively:

(3) Andrew is tall.

(4) False! Andrew is not tall. Mark is tall.

It seems plausible to say that they are disagreeing about *what it is to be* “tall”, or what *counts as* “tall”. If these sorts of disputes license the desired semantic assessments, why not just apply the same, or a roughly similar, model to normative disagreements?

As I see it, the main difficulty for this approach can be presented as a dilemma: either the alleged metalinguistic disputes license *metalinguistic negation* or they don’t. A metalinguistic negation (contrast: descriptive negation) is Horn’s (1989) name for a device by which someone can object to the non-semantic content of a previous utterance (implicature, presupposition, connotation, etc.). If metalinguistic *negotiation* occurred through metalinguistic *negation*, the metalinguistic strategy would deliver a good solution. After all, the latter is an independently attested linguistic mechanism and if normative disagreements could be analysed in that way, the “meta-analysis” would have a good deal of support. However, there are reasons to be suspicious of this move, and indeed Sundell and Plunkett (2013, pp. 32-35) acknowledge that the linguistic denial exhibited by normative disputes fail the classical tests for metalinguistic negation.²²

²²There are three tests that originate with Horn (1989). (i) The first is that metalinguistic negation doesn’t incorporate morphologically (“I don’t believe it, I know it!” is felicitous, but “I disbelieve it, I know it” is not); (ii) the second is that metalinguistic negation involves “the ‘contradiction’ intonation contour (a final rise within the negative clause), followed by a correction clause, and contrastive stress on the offending item and its replacement” (Carston (1996)) (e.g. “I don’t BELIEVE it, I KNOW it”, where the small capitals signal the emphasis); (iii) the third is that metalinguistic negation doesn’t license so-called negative polarity items such as “any”, “at all”, “ever” (“Mia didn’t eat some of the cookies—she ate all of them” is felicitous, but “Mia didn’t eat any of the cookies—she ate all of them” is infelicitous).

We're left with the second horn of the dilemma. On Sundell and Plunkett's proposal, the denials featured by normative disputes are ordinary descriptive ones which "pattern logically with denials of relevance implicatures" (p. 34). There is a natural worry that arises at this point, and it is this. If there is no linguistic evidence for taking these disputes to be metalinguistic, rather than descriptive, why should we construe them as such? Why not view all of them as ordinary disagreements where the two parties have a common property in mind and make conflicting claims about the extension of that property? Of course, one might reply by saying: because this is incompatible with contextualism. But this looks unsatisfying until a more detailed account is provided that explains why, given the conventional meaning of normative sentences, two syntactically contradictory utterances trigger the relevant metalinguistic negotiation.

Silk (forthcoming) offers a new contextualist framework for deontic modals, which he calls "Discourse Contextualism" that purports to address precisely this worry. On this view, deontic modals are semantically associated with a "discourse-level contextual parameter" which represents the norms endorsed in the conversation. This semantics generates constraints on the interpretation of deontic modal utterances, and disagreement is located in the incompatible assumptions about the value of the contextual parameter, that are carried by two contradictory utterances. This is a new and, in my opinion, very promising account. [...]

2.3 Expressivism

2.3.1 Introducing Expressivism

The basic non-cognitivist idea is that the conventional meaning of moral sentences like "eating fish is wrong" is more similar to that of exclamations ("Boo X!") and imperatives ("Don't do X!") than to that of descriptive statements, indexical or otherwise. Non-cognitivists think that terms like "good", "right" and "ought" are quite unlike ordinary descriptive terms such as "sunny", "young", or "being to the right of" in that they display a distinctive *practical*, *directive* or more broadly *non-cognitive* role, which makes them akin

to terms like “hooray”, “boo” or “ouch”. Expressivism—the view articulated, in different forms, by Blackburn (1993), and Gibbard (1990) and (2003a)—is the most popular, if not successful, non-cognitivist proposal. What distinguishes expressivism from other non-cognitivist analyses—like the emotivism of Ayer (1952), or the prescriptivism of Hare (1952)—is the fact that the former is first and foremost a view about moral *thought*—“it gives an account of moral language *by* giving an account of moral thought” (Schroeder (2008a), p. 4). In particular, the key expressivist thesis is that moral sentences express desire-like mental states in exactly the same way in which descriptive sentences express belief-like mental states.²³

It is sometimes said that expressivists think that normative sentences, like exclamations and commands, are not truth-evaluable or are not strictly speaking truth-evaluable. This is sloppy. A key feature of contemporary expressivism is deflationism about truth, i.e. the claim that the notion of truth is exhausted in terms of some biconditional schema²⁴, that the truth predicate doesn’t play any explanatory role, etc. This allows expressivists to say that normative sentences are truth-apt, and thus to accommodate everyday ascriptions of truth and falsity to normative statements.²⁵

In light of this, expressivism is best characterized as the view according to which the semantic properties of a normative sentence φ are to be explained in terms of properties of the *mental attitudes* conventionally expressed by utterances of φ , rather than in terms of properties of the *content/truth-conditions* of φ (cf. Silk (2015), p. 2; Schroeder

²³I will mainly speak to *pure* versions of expressivism here. According to *hybrid* expressivists, an utterance of a normative sentence expresses both a belief-like state and a desire-like state. See Schroeder (2009) for an overview. Moreover, by “expressivism” I will refer to *semantic* expressivism, rather than to *metasemantic* expressivism *à la* Ridge (2014) (see above), or to *pragmatic* expressivism *à la* Yalcin (2011) and (2012). According to Yalcin, expressivism is best understood as “a pragmatic thesis about natural language”, in particular as the hypothesis that “the common ground, beyond reflecting what is mutually presupposed about matters of fact, admits of something like plan-like structure, and that some sentences—the normative or plan-laden kind—serve particularly to manipulate this structure, ruling admissible plans in and out” (Yalcin (2012), p. 12).

²⁴This can be the disquotational schema “‘*s*’ is true iff *s*” (disquotationalism) or the equivalence schema “ $\langle p \rangle$ is true iff *p*” (minimalism).

²⁵It is important not to confuse the claim that normative sentences are *apt for minimal truth*, with the view called *minimalism about truth-aptitude*. There is a long-standing debate concerning the relation between *minimalism about truth* and *minimalism about truth-aptitude*, which revolves around the question: does the first entail the second? Boghossian (1990) says “yes”, Jackson et al. (1994) say “no”. However, everyone accepts that if *p* can be minimally true, then *p* is apt for minimal truth. This is all we need in this context.

(2008a)).²⁶ In particular, expressivists believe that sincere assertions of moral sentences express desire-like mental states instead of (representational) beliefs, and that the essential function of these assertions is to express desires, give advice or recommend, rather than to transfer information about the world. What is exactly the nature and structure of these desire-like mental states? Gibbard thinks of them as states of *norm acceptance*, or more recently as *planning states* (“hyperplans”); Blackburn appeals to states of *approval/disapproval*; Schroeder (2008) talks of an attitude of *being for*. In what follows I will sometimes use the term “desire-like state” as an umbrella term to cover the various alternative specifications of the relevant conative attitude.

2.3.2 Virtues and Vices

Expressivism has several theoretical virtues. An especially important one is that it neatly explains why, in general, speakers are motivated to act in accordance with their moral judgments. This falls out naturally from the expressivist hypothesis that to judge that an action is morally right *just is* to have a certain desire, and that this desire is inherently motivating. This thesis, i.e. the thesis that there is a necessary and internal connection between making a moral judgment and having some motivation to act accordingly, goes under the name of *motivational judgment internalism*. As far as I’m aware, all expressivists accept internalism—indeed, expressivism is often billed as the unique view that vindicates the truth of internalism.

But internalism, in the *unconditional* version just stated, is implausibly strong. Recall our discussion of cold and reverse uses: amoralists and immoralists may be psychologically puzzling figures, but they are not conceptually impossible. Expressivists have responded by adopting a *conditional* form of internalism, i.e. one that only applies if some “normal conditions” obtain, in particular if the moral agent is psychologically normal, practically rational, or morally perceptive (cf. Björklund et al. (2012)).

How is this reflected in the expressivist *semantics*? One might say that cold and

²⁶This seems to fit Gibbard’s own characterization: “[by] the term ‘expressivism’ I mean to cover any account of meanings that follows this indirect path: to explain the meaning of a term, explain what states of mind the term is used to express” (2003, pp. 6-7).

reverse uses require a different, non-expressivist, semantics—expressivism would still capture the allegedly “central” or “normal” uses. But this bifurcation is *ad hoc*, and fits poorly with the search for unification, which is often seen as a guiding principle in theory-choice. A more plausible hypothesis consists in complicating or weakening the kind of attitude (invariantly) expressed, e.g. by saying that it is a dispositional desire that needn’t always be occurrent (cf. Strandberg (2012), although he criticizes the view), or a second-order desire that may be overridden by a first-order one (as in Lewis (1989)), or that desire-like states are “multi-track dispositions” (involving motivational dispositions, but also non-motivating feelings) that may or may not manifest themselves in motivational states (see Toppinen (2015)). If any of these proposals can be made to work, the expressivist could be off the hook. But notice that this involves a switch to a somewhat different version of expressivism from the one we were accustomed to. If the foregoing considerations are correct, the conative attitudes expressed by moral terms are action-guiding and motivating in a subtler and weaker way than originally assumed. In this respect, many of the standard models of attitude-expressing terms, such as “Hooray” or “chink”, seem ill-suited to characterize the nature of the attitudes involved in moral judgments and assertions.²⁷ If anything, the “parity thesis” (Schroeder (2008b)) according to which normative sentences express desires in exactly the same way in which descriptive sentences express beliefs, seems highly controversial.²⁸

Perhaps the most pressing challenge for expressivists is the so-called “Frege-Geach Problem”, so much so that one of its more acute interpreters claims that expressivism is “a very unpromising hypothesis about the semantics of natural languages” (Schroeder (2008a), p. x). In broad strokes, the challenge is to make good on the idea that the meaning of a normative term consists in the attitude it conventionally expresses—and in this sense “wrong” behaves like “Boo!”—while accounting for the fact that, on the face

²⁷One might say that this is a problem in moral psychology, and as such is everyone’s problem. More precisely, it is a problem for any view on which some attitude or other is conventionally attached to the meaning of moral terms (or of some uses thereof—recall the internal/external distinction drawn in the discussion of contextualism). I don’t disagree with this. For example, Copp (2001) is not a pure expressivist, but Finlay (2005) has criticized his “conventional implicature” account of attitudinal content for making wrong predictions about (what I called) cold uses.

²⁸See Woods (2014) for an argument to the effect that the parity thesis is false.

of it, normative sentences like “eating fish is wrong” have the same syntactic, semantic, and logical properties as descriptive declarative sentences, like “eating fish is a source of omega-3 oils”.²⁹ Quickly, the problem is this. Consider:

- (5) If eating fish is wrong, then cooking fish is wrong.

If Jim sincerely utters the sentence: “eating fish is wrong” (or “cooking fish is wrong”) we learn that he disapproves of eating fish (or of cooking fish). This is not the case with (5). In asserting (5), Jim is not expressing his con-attitude towards any particular action (or action-type). So, what’s the meaning of “wrong” when it occurs in embedded contexts? Plausibly, “wrong” does not change its meaning as we move from freestanding to embedded contexts. After all, everyone agrees that moral *modus ponens* arguments—e.g. If eating fish is wrong, then cooking fish is wrong; eating fish is wrong; hence, cooking fish is wrong—are valid. But arguments that equivocate are in general not valid. For example, consider the following argument: (a) Mark is a fox; (b) if Mark is a fox, then Mark has four legs; (c) Mark has four legs. (a) is only plausible if “fox” is read metaphorically as meaning “a sly person”, whereas (b) is only plausible if “fox” literally refers to the animal. But this means that assuming that “fox” means something different in (a) and (b), the inference from (a) and (b) to (c) doesn’t have the *modus ponens* form (p ; if p , then q ; q).

The most promising solution to this and related problems is, arguably, the one offered by Gibbard. In a recent defense of expressivism, Silk (2014, pp. 6-7) nicely describes Gibbard’s project as follows:

The crucial contribution of Gibbard’s account, on my view, is that it starts with an existing framework for doing formal semantics—possible worlds semantics—and then systematically provides that framework with an expressivist interpretation. The compositional semantics proceeds roughly as usual; the abstract objects that Gibbard uses as semantic values—sets of world-hyperplan pairs—function like sets of possible worlds. What makes Gibbard’s

²⁹To be a little more precise, both normative and descriptive sentences can be embedded in complex linguistic environments, either truth-functional or non-truth-functional (as in: “eating fish is wrong and helping the poor is good”, “if eating fish is wrong, then cooking fish is wrong” or in: “Jim believes that eating fish is wrong”); both can figure in so called “mixed contexts” (such as: “either eating fish is wrong, or eating fish is a source of omega-3 oils”); both can be employed in valid inferences; and both are such that the meanings of the complex sentences thereby obtained is a function of the meaning of their parts and the way they are combined.

use of the framework distinctively “expressivist” is the interpretation he gives for it.

So, by saying that the semantic values of normative sentences are sets of world-system of norms pairs (Gibbard 1990) or sets of world-hyperplan pairs (Gibbard 2003), an expressivist can exploit the resources of this extended possible-world semantics to explain compositionality and logical validity. However, as the quotation makes evident, there is something that is distinctive of an expressivist semantics *vis-à-vis* a descriptivist, truth-conditional one: namely, that given the order of explanation that expressivists urge, the semantic properties of sentences, and the logical relations among sentences, are ultimately to be explained in terms of properties of, and relations among, the mental states associated with the relevant sentences, and this point should be reflected in the expressivist semantic analysis.

With the expressivist project so understood, some have worried that Gibbard’s solution to the Frege-Geach Problem is not satisfactory.³⁰ In this respect, a fully descriptivist semantics, like relativism, may be superior to expressivism, insofar as it shares its benefits and lacks some of its problematic commitments. As anticipated, relativists themselves employ more fine-grained intensions than just sets of worlds. Unlike expressivists, though, they would be free from the burden of an “attitude semantics”³¹ (The relationship between relativism and expressivism will be further examined in the next section).

³⁰See Schroeder (2008a). For some more recent solutions to the Frege-Geach Problem, see Silk (2015) and Baker and Woods (2015)

³¹My understanding of expressivism mainly follows Schroeder’s characterization in his (2008a) and (2015). Schroeder thinks that “an expressivist semantics can’t work by simply applying the same sorts of techniques to its ultimate semantic values as a truth-conditional semantics can” (2015, p. 229). The core of the tension is that according to expressivism the ultimate semantic values of sentences are belief-states. However, belief-states, unlike the propositional contents assigned by truth-conditional semantics, do not obey the principles of complementation, intersection, and union (e.g. believing that $\neg p$ is not the same as not believing that p). And while one can “apply those principles *indirectly*” to the propositional contents rather than to the belief-states, this move would fit poorly with the expressivist paradigm. This is what motivates Schroeder’s “biforcated attitude semantics” as the most natural semantics for expressivists. There are other ways of understanding expressivism, and of distinguishing it from relativism. For example, as mentioned above (cf. fn. x), Yalcin (2011) suggests that expressivists should characterize their view in contrast with relativism by denying that “the notion of truth at a context is well-defined for the sentences of his target discourse”. The difference would thus emerge in the postsemantics that the expressivist, unlike the relativist, doesn’t provide.

2.4 Relativism

2.4.1 Introducing Relativism

According to Relativists, an utterance of “eating fish is wrong” expresses the same proposition regardless of the context of utterance, i.e. regardless of who uses the sentence, or when they use it. However, this proposition can be true relative to Jim’s standard and false relative to Mia’s standard. The idea is that the truth of a normative proposition is relative to a standard or perspective in pretty much the same way in which the truth of a contingent proposition, like *that eating fish is a source of omega-3 oils*, is relative to a possible world (it is true in the actual world, but may be false in a different possible world). Formally, this result is achieved by adding a new coordinate/parameter, i.e. a parameter for a normative standard, in the Kaplanian circumstance of evaluation/index.³²

How is the truth or correctness of a use of a normative sentence determined, according to the relativist? There are two possibilities. The first consists in saying that the context in which a sentence is used determines the relevant circumstance of evaluation and hence the relevant standard with respect to which the sentence is to be assessed for truth and falsity. This would be analogous to what happens with standard relativity to possible worlds (cf. Chapter 1, §1.5). On this model, the truth-value of, e.g., Jim’s utterance of “eating fish is not wrong”, is fixed once and for all by the standard operative in Jim’s context. In a slogan: *propositional truth* is relative, but *utterance truth* is absolute. We can call this view *Use Relativism*.

2.4.2 Virtues and Vices

Relativists can offer a good explanation of disagreement, in the sense explicated above (cf. §2.2 above). If A’s utterance of ϕ expresses the proposition that p and B’s utterance of $\neg\phi$ expresses its negation $\neg p$, the two sentences are inconsistent in that A/B cannot rationally accept what B/A said without changing her mind (cf. Kölbel (2004), pp. 304-305). In MacFarlane’s (2007) vocabulary, A’s and B’s mental states are non-cotenable.

³²In other proposals, by adding a *context of assessment* to the classical double-index framework. See below.

A more language-oriented characterization of disagreement so understood might be to say that A and B propose to update the context, or to change the common ground, in incompatible ways.³³

Moreover, according to MacFarlane, assessment-relativism is uniquely suited to explain why two assertions of contradictory moral propositions—the propositions expressed by “eating fish is wrong” and “eating fish is not wrong”—, unlike two assertions of contradictory temporal propositions—the proposition expressed by “it is noon” and “it is not noon”, which, assuming that there are tensed propositions, are contradictory—, are such that we think they cannot both be correct, or jointly accurate. We think that an utterance of “it is noon” (at noon) and one of “it is not noon” (at midnight) can both be correct. This is explained by saying that their correctness depends on the context in which they are used. However, it seems to be part of our ordinary concept of normative disagreement that if two persons disagree, there is a sense in which they cannot both be correct (at least, this what many authors, MacFarlane included, assume). The challenge is to capture this in a way which is compatible with variabilism. Assessment-relativism can formally represent this by saying that the correctness of moral utterances turns on the assessor’s standard, and assuming he has a consistent standard, no two contradictory moral utterances can be true relative to that standard. So, relativists can explain the intuition that the two subjects cannot both be right without opting for variabilism or for the error theory, but by relativizing the notions of assertoric truth and correctness. This move also allows us to explain why we can retract or “take back” a prior speech-act in the light of a change in our views, since it is the assessor’s current standard that is relevant to truth-falsity evaluations.

However, it is precisely this innovation that makes relativism vulnerable to what is sometimes called “Evans’ challenge”.³⁴ The challenge consists in giving an account of what “we aim at, or take the others to be aiming at”, in asserting a proposition, if the correctness or incorrectness of that speech-act can’t be subject to a “once-and-for all

³³See Egan (2010) for suggestions in this direction.

³⁴Evans raised the objection against Prior’s time-neutral propositions in his “Does Tense Logic Rest on a Mistake?”.

assessment” (Prior 1985, p. x). Relativists have offered various solutions. For instance, MacFarlane (2014) proposes a commitment-based theory of assertions, according to which to assert that p is to undertake certain commitments or obligations with respect to p , for example to vindicate it if challenged, to withdraw it if shown to be untrue, and to be accountable or responsible if others rely on it. MacFarlane offers a relativist version of these three norms (which I provide in footnote).³⁵

Some have argued that this doesn’t solve the basic problem, because it introduces commitments that “rational and earnest speakers” cannot possibly undertake (Marques 2014). Briefly, the argument is this. There are cases in which speakers can felicitously resist to retract a certain claim. But if retraction was obligatory, these speakers would be perceived as violating some norm, perhaps as being irrational or insincere. So, retraction cannot be obligatory.

In conclusion, we could perhaps say that “Evans’ challenge” is for the Relativist what the “Frege-Geach problem” is for the Expressivist. It is an explanatory task, or perhaps burden, which any view that introduces theoretical innovations typically faces.

2.4.2.1 A Dilemma for Relative Propositions

There is a second worry that I want to raise, and which targets more specifically the idea that *propositional* truth is relative. Since the argument requires some elaboration, I will dedicate a separate section to this issue.

The problem can be stated as a dilemma for the metaphysics of relative propositions, where by “relative propositions” I (stipulatively) refer to propositions the truth-values of which may vary along a subjective parameter, like a moral standard. If, as I think, none of its horns is particularly appealing, this might cast some doubts on the overall

³⁵(**W**)*: In asserting that p at C1, one commits oneself to withdrawing the assertion (in any future context C2) if p is shown to be untrue relative to context of use C1 and context of assessment C2.

(**J**)*: In asserting that p at C1, one commits oneself to justifying the assertion when it is appropriately challenged. To justify the assertion in a context C2 is to provide grounds for the truth of p relative to context of use C1 and context of assessment C2.

(**R**)* In asserting that p at C1, one commits oneself to accepting responsibility (at any future context C2) if, on the basis of this assertion someone else takes p to be true (relative to context of use C1 and context of assessment C2) and it proves to be untrue (relative to C1 and C2) (cf. MacFarlane 2005, p. 337).

attractiveness of the relativist package. The dilemma is this: either relative propositions represent the world/express facts (call this view *factualism*) or they don't (call this view *non-factualism*).³⁶ If they do, some more work is needed to make the underlying required metaphysics plausible. If they don't, some significant differences between relativism and other non-truth-conditional semantics seem to evaporate. It seems safe to say that factualism and non-factualism exhaust the logical space of possibilities, and so the dilemma looks like a genuine one.

One might think that relativism is a view in semantics, and that semanticists could simply stay neutral on metaphysical issues. Brogaard (2008), for instance, hints at the question of whether the relativist's functions from standards to extensions "correspond with any (sparse) properties", but leaves the issue open, pointing out that "it is not the job of a semantic theory to say whether there exists a certain kind of property" (p. 392). This is a fair and uncontroversial remark, as far as it goes. But it does not touch the present worry. Plausibly, if factualism and non-factualism fully cover the range of possibilities, the relativist is committed to the truth of the disjunction of factualism and non-factualism. In light of this, I think that the *divide et impera* move has no bite here.

Most of what I'm going to say here assumes that there is a distinction to draw between the "semanticist's properties/propositions"—properties/propositions in the sense of *contents* or *intensions* of predicates/sentences, i.e. functions from indices to extensions—and the "metaphysician's properties/propositions"—i.e. properties/propositions understood as *worldly features*, or *truth-makers* (cf. Chapter 1, §1.5). To avoid confusions, especially when predicates are involved, I'll use "intensions" or "contents" for the semanticist's notion, and "properties" or "truth-makers" for the metaphysician's notion. Let us turn to spell out and discuss the two horns of the dilemma.

There are at least two factualist options for the relativist. I will articulate them by using a distinction that Spencer (2014) usefully draws. Spencer claims that there are two metaphysically distinct ways of thinking about the truth-value relativity of a certain

³⁶If you're squeamish about *facts*, you can recast the dilemma in terms of *truth-makers* or in terms of *monadic truth/truth simpliciter*. Nothing should hinge on this for my purposes.

proposition: *relationalism* and *variabilism**.³⁷

The *relationalist* thinks that the proposition *that eating fish is wrong* may change its truth-value across standards because of a “relativity in truth-making”. More precisely, the relationalist holds that “different properties are alethically relevant to the same proposition” relative to different standards. These properties will be, for instance, the property of *being wrong according to the vegan standard* and the property of *being wrong according to the omnivore standard*. I take it that the most straightforward way of cashing out relationalism and the talk of “alethic relevance” goes as follows. Propositions represent the world *perspectivally*, as it were. Representation is not a binary relation between two entities (the *representans* and the *representandum*) Rather, representation is a *ternary* relation between a proposition, a world and a standard. The truth-value relativity of *p* is thus explained in terms of a relativity in the expression relation between the proposition and the fact it expresses.

The *variabilist**, on the other hand, thinks that “the same property is alethically relevant to the proposition” irrespective of which standard we choose. However, this property has a variable extension, i.e. an extension that varies across standards. In other words, the *variabilist**, unlike the relationalist, thinks that a property may be relative to a parameter in a way which doesn’t involve increasing the adicity of the property by adding an extra argument place.

It is important to stress that Spencer intends the relationalism/*variabilism** distinction as a distinction *within* relativism. In principle, a relativist semantics can be combined with either metaphysical view. So, for instance, he points out that “there might be good reason to adopt truth relativism about epistemic modals [...] even though there is very little reason to believe in the unary property of might-ness” (*ibid.*, p. 14).³⁸

Wright (2008) argues that the prospects for relativism cum *relationalism* are grim. The relativist needs the representational content to be invariant, and no example of

³⁷The starring is mine and is meant to signal that this thesis has nothing to do with what I’ve been calling “variabilism” so far.

³⁸More generally, Spencer thinks that the metaphysical distinction between relationalism and *variabilism** is orthogonal to the one between contextualism and relativism in semantics. So, in principle all four combinations of views are available (although the contextualism-*variabilism** pairing is admittedly the most awkward) (cf. pp. 16- 17).

perspectival representation seems to deliver that. So, I think of proof is on the relationalist relativist to explain how relative propositions can represent in a perspectival way. We're left with variabilism*. On the face of it, this looks like a more promising view. However, I will now argue that variabilism* brings with itself some metaphysical commitments that may be hard to swallow.

On the face of it, the variabilist* proposal requires that the intension of the predicate "wrong" stand for a *non-natural* property. For suppose it stands instead for a property that is monadic and yet natural. Then it is hard to see how two contradictory utterances involving "wrong" could both be correct/true/faultless. Barring run-of-the-mill vagueness/indeterminacy in the, possibly highly conjunctive or disjunctive, descriptive predicate to which the moral predicate "wrong" is reduced (for the reductive naturalist), or in the irreducible moral predicate "wrong" (for the non-reductive naturalist), there will be a fact of the matter as to whether a certain action-type, say, *eating fish*, possesses or fails to possess that property.

Now, ask yourself if there is such a thing as the monadic and non-natural property of being wrong. Answer "no", and you have an empty predicate, relativizing the extension of which doesn't take us anywhere. Answer "yes" and you fall to the usual complaints that these properties would be *queer*, "utterly different from anything else in the universe" (Mackie 1977, p. 38). Extensional shifts would hardly improve their ontological status. Compare perhaps a relativist treatment for the predicate "being a god". The view says that the semantic content of "Zeus is a god" is a proposition that may be true according to a religion R1, and false relative to a religion R2. But, of course, if the property *being a god* is not in a good standing, relativizing its extension doesn't make it any better. Polytheism is not a good solution for a god sceptic.

Perhaps, the relativist could respond by saying that standards are not like religions. Standards are more like possible worlds. A possible world is constituted/composed by the facts that obtain in it. A moral standard may likewise be constituted/composed by the moral facts that obtain in it. This analogy needs some spelling out. In particular, what determines the identity of the facts that make up a standard? Arguably, we cannot say

that these moral facts are fixed independently of any subjective factor. For if this was the case how could a standard have any influence on the truth-value of a moral proposition? In other words, this would take us back to the “god” case. The alternative is that the identity and obtaining of these facts, for instance of the fact that eating fish is wrong, partially depends on a given subject’s desires, preferences, attitudes, and so on. If we think of the truth-making relation along these lines, then we can start to see how a moral proposition might be made true by a standard in a way that parallels the one in which a factual proposition is made true by a world.

I’m aware of three attempts at fine-tuning this suggestion: Wright’s (2008) “truth-conferring model” of relativism, Einheuser’s (2008) “factual relativism” and Rovane’s (2012) “multimundialism”. Some quotations will prove useful.

We might for example permit the actual world at Williamson—that is the actual world as reflected in Timothy Williamson’s gustatory standards—to exist simultaneously with the actual world at Wright, that is, the actual world as reflected in Crispin Wright’s gustatory standards. The proposition that stewed rhubarb is delicious can then be true at the one aesthetic location, so to say, and untrue at the other, just as the proposition about the sea fight can be true at t_2 but untrue at t_1 . [...] The relativism surfaces [...] in the thought that there is no single actual world but a plurality of them. [...] It abandons the idea of a single, comprehensive Tractarian world—a totality of all facts. (Wright 2008, p. 172)³⁹

The actual substratum from my perspective, represented by $s@$, pI —with $s@$ being the actual substratum and pI being my perspective—is a world in which the general taste facts are determined by my sensibilities. It differs from, say, the actual substratum from Moira’s perspective, $s@$, pM : In $s@$, pI Vegemite tastes awful while in $s@$, pM it tastes great. Different possible worlds with the same substratum coincide in their objective facts, but differ in their subjective (general) taste facts, as these are induced by different perspectives. (Einheuser 2008, p. 191)

Relativists [should] affirm instead that some truth-value bearers do not stand in logical relations to one another, that there are many noncomprehensive bodies of truths that cannot be conjoined, that there are many worlds rather

³⁹*Nota Bene*: Wright articulates, but doesn’t endorse the view.

than one. In other words, they [should] affirm Multimundialism. (Rovane 2013, p. 91)

While there may be differences in the details of these three views, they all put forward the *many-worlds* idea. There are, I think, two main challenges for any view of this sort.

The first consists in coherently fleshing out the thought that there is a plurality of subjective worlds. To put it in Wright's terms, what does it mean to say that the actual world at Williamson and the actual world at Wright "exist simultaneously"? One might think of these two worlds as composing a single overarching world, perhaps in the same way in which Europe and Asia are disjoint locations of a single all-inclusive spatial world. But this can't be the right way to interpret the proposal. Taking this route would involve adopting a relationalist metaphysics—i.e., there are two different yet compatible facts, the fact that standard S1 forbids eating fish, and the fact that standard S2 permits eating fish—or admitting the existence of contradictory facts—i.e. the fact that eating fish is forbidden and the fact that eating fish is not forbidden. How should we then understand the many-worlds idea, if it doesn't involve a higher-order larger world, and the postulation of contradictory facts?

In my view, this question remains without a satisfactory answer in the aforementioned accounts. This is not a detail we can overlook. For the multimundialist package, if taken seriously, requires a huge departure from the standard way of thinking about reality or about what is the case. After all, on a natural conception, the question: "what is the case?" does have a unique answer. What if the multimundialist pulls back and say: "look, all I'm saying is that there is no fact of the matter as to whether p expresses/represents a fact or not. p may represent a fact for A, but not for B". My reply is that this move is either a restatement of the many world-idea, or it amounts to a (perhaps sophisticated) version of non-factualism.

The second worry concerns the metaphysical extravagance of subjective facts. Consider Einheuser's metaphysical picture. A possible world is a substratum-perspective pair $\langle s, p \rangle$; in other worlds, it is the result of combining a subjective perspective with an objective substratum. In the moral case, the substratum might determine all the physical

properties of a certain action (or the physical properties plus the psychological properties, if you think the latter do not supervene on the former). For instance, these properties may be: that eating fish produces pain in fish, that eating fish is a source of omega-3 oils, that eating fish is not necessary for survival, that eating fish is disapproved of by Mia and not disapproved of by Jim, and so on. While we all share the objective substratum, each of us “induces” his or her own perspective on it. Moral propositions represent such subjective-perspectival facts, and different people inhabit different worlds, as configured or “distorted” by their own perspectives.

If taken at face-value, or in a heavyweight fashion, this picture seems to be in violation of the widely shared thesis that moral properties supervene on physical properties. Take a substratum *s* as our supervenience base. The supervenience thesis, couched in the new vocabulary, says that:

(S): For all substrata *s* and *s**, if the distribution of physical properties at *s* and *s** is exactly alike, then the distribution of moral properties at *s* and *s** is also exactly alike.⁴⁰

But that’s precisely what Einheuser’s factual relativism, literally understood, denies. This is metaphysically puzzling. Moral facts seem to partially float free of the physical facts. Their queerness might be different from the one targeted by Mackie, but it’s not clear that it is less worrying. On the contrary, there is reason to suspect that it is *more* worrying: not only moral properties are not reducible to physical properties; they don’t even supervene on them.

Like I said, these consequences only follow if we take the appeal to subjective worlds *literally*. Perhaps, charity suggests a different reading. How else could we interpret the framework? Commenting on Einheuser’s proposal, and relating it to his own version of relativism, Kölbel (2009, fn. 13) says:

Einheuser (2008) uses a terminological variant: what I call “possible world”, she calls a “substratum”, and what she calls a “possible world”, I call an or-

⁴⁰This is a very approximate formulation of supervenience, for reasons illustrated, among others, by Jackson (1998). But I think we can safely ignore these niceties for our purposes.

dered pair of a possible world and a standard of taste. Einheuser seems to think that this variation makes a metaphysical difference [...]. I doubt the metaphysical significance of this terminological move: it merely sounds more exciting to say that we don't share the same world than to say that we don't share a standard of taste. But then I may not be in disagreement with Einheuser, for her conception of metaphysics or ontology seems to be pretty light: she claims that "Facts and meanings are theoretical concepts introduced to systematically account for our linguistic behaviour and cognitive interaction with the world." (p. 189).

If Kölbel is right,⁴¹ then "factual relativism" is a bit of a misnomer for Einheuser's view. For at least on the present taxonomy, her proposal comes out as a variant of non-factualism.⁴²

Let's wrap up our discussion of the first horn of the dilemma. The most plausible factualist metaphysics for relativism goes along the lines of Spencer's variabilism*. The few proposals advanced in the literature are underdeveloped in ways that make it difficult to work out their distinctive, if any, metaphysical commitments. But if the above considerations are on the right track, this much seems true. If the appeal to subjective worlds is to be taken in a lightweight fashion, then this version of relativism simply collapses into a form of non-factualism. If, on the other hand, we take the proposal in a heavyweight fashion, then we encounter a number of worries about the nature of these facts. For one thing, it is unclear how the threat of contradictory facts can be escaped; for another, even setting that aside, the problem remains that we are saddled with facts that are far from being metaphysically conservative.

Do relationalism and variabilism* exhaust the factualist routes for the relativist?

⁴¹In her (2006), Einheuser seems to confirm that she has in mind the lightweight interpretation. She models a possible world as a substratum-carving pair (a carving being roughly equivalent to she calls a perspective in her later paper), but in a footnote she says: "It is possible to proceed by equating conventionalist possible worlds with substrata, and letting propositions [...] be true or false at a world relative to a carving. Analogues of all of the concepts I introduce in this paper can be introduced for this construal of conventionalist possible worlds and the accompanying relative notion of truth at a world. Therefore, nothing of substance hangs on the particular choice of representational primitives made here." (p. 481, fn.7).

⁴²See Marconi (2014) for some other compelling considerations against the appeal to subjective worlds.

I'm not sure. But it's hard to find other minimally plausible alternatives. For instance, consider the view that relative propositions are representational in the sense that they purport to represent monadic facts (e.g. the fact that eating fish is wrong), but as it turns out the world only contains relational facts (e.g. the fact that eating fish is wrong relative to the vegan code). This hypothesis would entail that ordinary speakers are mistaken in their moral ascriptions, a prediction that relativists surely want to avoid.

In any case, I'm not claiming that there is no other plausible metaphysics for the factualist relativist. So one can think of the upshot of the above considerations as a challenge: if you want to be a factualist relativist, give me a plausible metaphysics of relative facts.

Most (perhaps all) relativists implicitly adopt some version of non-factualism. They think that propositions are not bound to be *representational contents*. Schroeder (2011) nicely illustrates how this result can be achieved while sticking to the idea that propositions play some explanatory or theoretical work (contrast: simply allowing for a deflationary accommodation of proposition talk⁴³). The idea is that there are two job descriptions for propositions, which correspond to two distinct theoretical roles. On the first conception, propositions are the semantic values of declarative sentences (in context), the objects of attitudes and speech-acts, and the primary bearers of truth and falsity. On the second conception, propositions carve up the world at its joints, and are associated with metaphysical commitments (p. 410). The propositions of our non-factualist relativist will meet the first but not the second job description: they needn't be anything more than some set-theoretic or function-theoretic object.⁴⁴

Relativists of this sort are not committed to any extra-ontology above and beyond the one accepted by the contextualist (or by any orthodox naturalist, for that matter). The world doesn't contain the monadic moral property of being wrong, only the descriptive properties of being wrong according to the vegan standard, being wrong according to

⁴³This deflationary route is followed, for instance, by Blackburn (1993). On this line we can say things like: "There is the proposition that eating fish is wrong" just because, and insofar as, the embedded sentence "eating fish is wrong" is syntactically well-formed and meaningful.

⁴⁴This, in turn, would be due to the fact that the semantic values of predicates are not construed as real properties, but rather as set-theoretic or function-theoretic objects.

the omnivore standard, and so on. In his non-factualist clothes, the relativist holds that what makes p true relative to an assessor A is not some relative fact about wrongness, but a broadly psychological fact about that assessor's preferences, values, etc.

Why should this non-factualist route be problematic at all? Let's start by considering an intuitive, perhaps naive, puzzle. If there is no such thing as the fact that p , for instance no such thing as the fact that eating fish is wrong, how can A 's judgment/assertion be true rather than merely indeterminate in truth-value, if not flat-out false? One natural reply is this: what grounds the truth of A 's judgment that p is the fact that p is true relative to A 's standard. But of course this just pushes the problem back a level: what does it mean to say that p is "true for A "? And what is the relationship between truth relative to a standard and monadic truth or truth *simpliciter*?

Two prominent relativists, Kölbel and MacFarlane, seem to agree that there is no straightforward reduction of ordinary propositional truth to truth relative to a standard. Let's use "trueO" for the everyday notion of truth and "trueS" for the semanticist notion of truth. Note that it is common ground and uncontroversial that the notion of truth used in semantics differs from the ordinary notion of truth, in that it applies to *sentences* (contrast: contents of speech-acts and attitudes) and it is *polyadic* (contrast: monadic) for sentential truth is typically relativized (at least) to a model, a variable assignment, a context, and an index. What is at issue is the presence or absence of analytic-reductive links between ordinary propositional truth and the various relativized notions used in semantics.⁴⁵

Kölbel (2008) considers two possible reductions of ordinary truth to truth relative to a standard—to say that (i) p is trueO iff p is trueS at "the privileged standard" (in analogy with the actual/privileged world) or to say that (ii) p is trueO iff p is trueS according to the standard of the speaker of the context. But he points out that (i) would make one of the parties to the disagreement faulty, and (ii) "would commit the relativist semanticist to an outright contradiction" (p. 249), insofar as p and not- p would both come out true

⁴⁵Like I said, in truth-conditional semantics "true" applies to *sentences* or *utterances*; a notion of truth for *propositions* is typically defined in terms of sentential truth. However, since we are interested in propositional truth, I will simplify and couch most of what follows directly in propositional terms.

(just imagine two contexts occupied by speakers with conflicting standards). Likewise, MacFarlane (2014) doesn't define a monadic truth-predicate for propositions in terms of the dyadic "true as assessed by X". The basic notion of his semantics is that of "true as used at c_1 and assessed from c_2 ".

However, there seems to be a difference between Kölbel's and MacFarlane's positions, one that may affect the overall evaluation of the second horn of the dilemma. I will briefly summarize what I take to be the main difference, and will then raise some concerns. Kölbel (2008) suggests that the theoretical notions of sentential truth, and propositional truth, are "not grounded in our pre-theoretic notion of truth, as the standard semanticist maintains, but rather [...] in a different pre-theoretical notion of correctness" (p. 250). The idea, as I understand it, is this. Rather than saying that p is true_O iff p is true_S according to the standard of the speaker A , we should say that it is correct for A to assert that p iff p is true_S according to A .

One natural worry is that the ordinary notion of truth seems to play no role in this new characterization. But if we cannot ground truth_S in our everyday notion of truth, why call it "truth" in the first place, and why call the resulting semantics a "truth-conditional" semantics? Kölbel (2001, 2008) offers two answers. The first is that there is a significant overlap between the two notions, because "in the objective range" correctness and truth_O coincide. For instance if we consider the proposition expressed by an utterance of "eating fish is a source of omega-3 oils" then the correctness of that utterance can be identified with the truth_O of the proposition expressed (Kölbel 2008, pp. 250-252). The second is slightly more complex, and the details need not detain us here, but the basic idea is that the disquotational properties of the ordinary truth-predicate serve as an expedient in the recursive machinery of the theory [of meaning]" (Kölbel 2001, pp. 633-635).

MacFarlane (2014), instead, seems to think that truth_S, i.e. the semanticist's notion of truth, is grounded in our pre-theoretic notion of truth. But how can this be if truth *simpliciter* cannot be identified with truth_S relative to some privileged assessor? MacFarlane entertains, but immediately rules out one popular strategy: that of "giving a definition of truth that makes its assessment-relativity plain" (p. 97), for instance a

definition of truth in terms of idealized justification (perhaps “superassertability”), or in terms of what is good to believe. His rationale for dismissing these options is that while they capture “the ‘relative’ part of ‘relative truth’”, they don’t capture “the ‘truth’ part” (p. 98).⁴⁶

But *what* should be captured exactly? As I understand his view, MacFarlane seems to buy into Dummett’s (1978) claim that a truth-conditional theory of meaning presupposes an independent grasp of the notion of truth, and is therefore incompatible with the idea that such notion is exhausted in terms of some biconditional schema or that the truth predicate doesn’t play any explanatory role. Roughly, Dummett’s argument is that if one accepts that the Tarskian T- sentences (e.g. “‘the snow is white’ is true iff the snow is white”) exhaust all there is to say about the concept of truth, then one cannot at the same time regard the T-sentences as meaning-giving, on pain of circularity. Since MacFarlane thinks that the T-sentences are meaning-giving, he must have in mind some non deflationary notion of truth. The most plausible conjecture is that he opts for a kind of primitivism about the concept of truth.⁴⁷ (He seems to confirm this speculation, when he cites Davidson’s take on truth on p. 98).⁴⁸

Of course, MacFarlane also holds that one can, and should, give an illuminating *explication* of the concept. Following in the footsteps of Dummett, MacFarlane suggests that the most promising strategy proceeds by connecting the truth predicate to the speech-act of assertion. In particular, his final proposal is to make sense of relative truth in terms of two pragmatic norms (the “Reflexive Truth Rule” and the “Retraction Rule”) which state the conditions under which it is correct to assert a relative proposition, and those under which one is required to retract a previous assertion. (So, MacFarlane’s response to the question: “why is the semanticist’s notion of relative truth our familiar

⁴⁶See also pp. 40-41 where MacFarlane sets out to explain why his notion of relative truth is “a relativized truth predicate, and not something else entirely”.

⁴⁷Asay (2013, p. 2) characterizes the primitivist doctrine as follows: “Primitivism is the view that truth is a fundamental concept. As such, it cannot be analyzed, defined, or reduced into concepts that are more fundamental.” Deflationists are not primitivists because they think that truth can be implicitly defined by the T-sentences.

⁴⁸On the other hand, I confess that I don’t fully understand how this primitivism can be reconciled with MacFarlane’s doubly relativized truth-predicate (“true as used at c1 and assessed from c2). Isn’t that *precisely* a defined truth-predicate, and as such something that we cannot grasp independently of the theory? [More work to do here].

notion of *truth* and not a stipulated new notion?” seems to be first (i) that our concept of truth, both the one used in semantics and the everyday concept, cannot be defined in more primitive terms; and second (ii) that the relativized truth-predicate used in semantics (and, given (i), *any* truth predicate) can be explicated in terms of the norms governing our assertoric practice.)

A brief recapitulation is in order. Kölbel and MacFarlane seem to suggest that all we need to understand in order to understand what it means for p to be true relative to A are the norms governing the correctness of assertions (and, for MacFarlane, retractions) of p . This is fine, as far as it goes. If the problem was one of intelligibility, “correctly assertable by A ” may give us a good way of making sense of “true for A ”.⁴⁹ However, there are a number of lingering worries.

The first is a *circularity* worry. One of the original motivations for relativism was its ability to capture the phenomenon of *faultless disagreement*, roughly the fact that for some p , it may be that A is entitled to assert that p and B is entitled to assert its negation, $\text{not-}p$. If this was the *explanandum*, the explication of relative truth in terms of correct assertability looks more like a restatement of the datum, than a genuine explanation. No doubt, this worry loses force if the relativist’s endeavour is characterized as merely *predictive* rather than explanatory in a more robust sense.

The second is a worry of explanatory *redundancy*. If the real explanatory work is not done by truth, but by some other notion, why not streamline the story and theorize directly in terms of that other notion? Namely, why not theorize directly in terms of the norms that govern the making of correct assertions/judgments? The worry, in other words, is that tying relative truth to correct assertability too closely exposes the relativist to the criticism, recently raised by Horwich, that his semantics “is, in substance, nothing but a use-theory in truth-theoretic clothing.”⁵⁰ (Horwich 2014, p. 749).

There are at least two things the relativist could say in reply. The first is that the no-

⁴⁹One might raise further worries about the normative notion of “correctness” invoked. But I will not fuss about that here, and will assume a relativist can give some dispositional characterization of this concept.

⁵⁰Where this amounts to “a theory in which the meanings of words are identified with the acceptance conditions of certain sentences containing them”.

tion of truth was never supposed to play a central explanatory role in a truth-conditional view of meaning, and so that Horwich's collapsing worry is not really problematic. Note that this reply is open to Kölbel, but is not really an option for MacFarlane, insofar as accepts the Dummett-Davidson thesis that truth does play an explanatory role in semantics. Note also that this reply might be problematic in any case. A wide number of philosophers, after all, have been convinced by Dummett's argument (briefly summarized above) that a truth-conditional semantics is incompatible with deflationism about truth, and some of them have opted for a use theory of meaning on the grounds that, unlike the truth-conditional approach, the former is compatible with deflationism about truth.⁵¹ So, Horwich's worry might ultimately be resisted, but it cannot, in my view, be dismissed out of hand.

A second, and perhaps most straightforward reply the relativist could help himself with is this. We need truth-conditions to explain the compositionality of meaning; non-truth-theoretic accounts of compositionality à la Horwich are not equally satisfying. But note that even if we granted this, it wouldn't follow that the notion of relative truth has explanatory value. Perhaps an improved and simplified truth-conditional semantics would only employ a notion of truth-at-a-world, but would be paired with the same pragmatic norms sketched by MacFarlane.

So, the non-factualist brand of relativism faces the challenge of explaining (i) why the true_S in "true_S relative to A" is properly called true_S rather than bla_S, where p is bla_S relative to A iff p is correctly assertable by A; (ii) why "true_S relative to A" is not explanatorily redundant, or dispensable, so that a simpler and more elegant theory would only appeal to a notion of assertoric correctness.

Let's take stock. I have claimed that relativists face a dilemma when it comes to specifying the standard-relativity of the truth of moral propositions. Either relative propositions represent some state of affairs or they don't. I have argued that the first horn of the dilemma has some unpalatable metaphysical consequences. On the other hand, the second horn is metaphysically conservative, but it is unclear whether the resulting

⁵¹Besides Horwich, a notable example is Field (1994).

view still deserves the name of truth-conditional semantics.

2.5 The Error Theory

2.5.1 Introducing the Error Theory

In its paradigmatic form, originally developed by Mackie (1977), the *moral error theory* consists of two claims: a *semantic* or *conceptual* claim and a *metaphysical* claim. The conceptual claim is that moral discourse is non-negotiably, or conceptually, committed to the existence of “objective values”. As Mackie puts it: “Ordinary moral judgments include a claim to objectivity, an assumption that there are objective values” (Mackie 1977, p. 35). The metaphysical claim is that “there are no objective values” (*ibid.*, p. 15). Three points need clarification.

First: what it is for a discourse *D* to be *non-negotiably* committed, or *conceptually* committed, to the truth of a certain proposition *p*, in this case to an existential claim? Roughly, the idea is that the commitment is subject-determining: if you drop the commitment, you simply change the subject. A morality which is not committed to the existence of objective values is, in fact, only a “schmorality” (Joyce 2008).⁵² So, there is no way in which we can correct the error while simultaneously retaining the theory.

Second: what it is for a discourse *D* to be *committed* to a certain claim *p*? There are various options: *p* can be *presupposed*, *entailed*, or perhaps *conventionally implicated* (in the Gricean sense). Mackie is a bit unclear on this point, but contemporary error theorists typically opt for the presupposition view (like Joyce 2001) or for the entailment view (like Olson 2014).

Third: how are we to understand the phrase “there are objective values?” In principle, the slogan can be given two distinct readings, which we may call the *lightweight* reading and the *heavyweight* reading. On the *lightweight* reading, the commitment is simply to the existence of a single uniquely adequate moral standard, the validity of which may be grounded in natural, and even mind-dependent, facts. So, for an instance an

⁵²“Any value system with the flawed element(s) extirpated simply wouldn’t deserve the name ‘morality’” (Joyce 2008, p. 52).

assertion of “eating fish is wrong” conceptually implies that eating fish is wrong according to the single uniquely correct moral standard. Smith (1994) and Jackson (1998), for instance, independently suggest that their views depend on something like the following presupposition: *there will be a single mature folk morality*. But this presupposition may well turn out to be false. Were this to happen, one could conclude that there are no moral properties (Smith 1994, pp. 187-189).⁵³ According to the *heavyweight* reading, the commitment is to the existence of categorical reasons, or objectively prescriptive facts, or irreducible (dis)favouring relations. So, for instance, an assertion of “eating fish is wrong” conceptually implies that there are categorical reasons not to eat fish, where a categorical reason is a reason that binds everyone irrespective of one’s desires, aims, ends, etc. (see, for instance, Joyce 2011 and Olson 2014)⁵⁴ But these facts—i.e. categorical reasons—would be entities “of a very strange sort, utterly different from anything else in the universe” (Mackie, p. 38). As Mackie famously put it, they would be queer, and so do not exist, at least in the actual world.⁵⁵

The heavyweight version is the one typically adopted by error theorists, so I will focus on it in my discussion. It may be important to note, though, that the conceptual commitments that the lightweight error theorist and the heavyweight error theorist ascribe to moral language and thought—call the two commitments *L* and *H*—are orthogonal.⁵⁶ For instance, as should have emerged from the previous remarks, one might think that moral language is committed to *L*, but not to *H*. Interestingly, the converse holds as well. One can imagine a community whose members refer to two different properties by using the term “morally wrong”. In the mouths of some members “morally wrong”

⁵³An alternative, endorsed by Jackson, is to say that moral terms pick out different properties in different folk moralities.

⁵⁴This feature, namely the fact that moral oughts entail *reasons for action* that are independent of people’s interests and ends is what, according to error theorists like Joyce and Olson, sets moral normativity apart from other varieties of normativity whose requirements have *some* categoricity, in that they *apply* to people irrespective of their interests and ends. Take the oughts of etiquette, games and grammar. From the point of view of etiquette, one ought not to speak with her mouth full, regardless of whether by so doing she would enjoy the food more. However, this doesn’t mean that one has reasons to act in accordance with the demands of etiquette independently of whether this would further one’s aims or not.

⁵⁵I will bracket the question of the modal status of the error theory, since it doesn’t matter much for our purposes.

⁵⁶This makes the two labels less than ideal, at least insofar as they suggest a relation of inclusion (e.g. if something is heavyweight, *a fortiori* it is lightweight). With this caveat, I will stick to them.

picks out the property of *failing to maximize happiness*, whereas in the mouths of some other members it picks out the property of *failing to comply with the Kantian imperative*. Moreover, this is transparent to language users in the same way in which it is transparent to them that “I” may refer to different persons when used by different speakers. However, “morally wrong” as used by this community conceptually implies *H*: saying “eating fish is wrong” conceptually implies that there are facts that counts against fishing. When the Millian says “eating fish is wrong” his claim conceptually implies that there are facts that count against performing actions that fail to maximize happiness; when the Kantian says “eating fish is wrong” his claim conceptually implies that there are facts that counts against performing actions that don’t follow from the universalizability maxim.

What conclusions should we draw if the conceptual claim is true, and the metaphysical claim is false? Error theorists conclude, plausibly enough, that all substantive moral claims are untrue, in that they fail to express a proposition, express a proposition that is neither true nor false, or express a proposition that is false.⁵⁷

2.5.2 Virtues and Vices

The error theory is often claimed to be the unique variabilist account that is able to accommodate the *authority* of moral discourse. It is not always clear what this means, but a natural gloss is that moral requirements have a quasi-magical “binding force” or “practical oomph” attached to them: they are supposed to provide reasons for action that exist and apply to agents irrespective of their desires, goals, and ends.

Not everyone agrees that this is a virtue, because it is controversial to what extent this element of authority is really constitutive of moral concepts. We can identify three styles of objections. (i) The first, pursued in Finlay (2008), says that there is linguistic *evidence against* the conceptual claim of the error theorist. (ii) The second, suggested by Lewis (2005) and also by Finlay (2008), has it that “there may be errors in morality as such but correction of them is not tantamount to the abandonment of morality. Morality

⁵⁷Which of these three options (or possible others) they choose turns on how they represent the defective commitment, e.g. as a presupposition or as an entailment, and on further issues in the semantics of non-denoting terms which need not detain us here.

may be erroneous, but not essentially erroneous” (Lewis 2005, p. 318). The thought is that the error, if there is any, is *semantically innocuous*, i.e. it is such that it doesn’t affect the conventional meaning—truth-conditional or otherwise—of moral statements. Or perhaps it is only *truth-conditionally innocuous*: while being conventionally attached to the meaning of moral statements, it doesn’t affect their truth-conditions.⁵⁸

(iii) A third objection, which builds again on a Lewisian suggestion, consists in saying that the linguistic evidence doesn’t decide whether the conceptual claim is true or not: “the content of morality is sufficiently ill-defined that we cannot show that any errors are errors of morality as such, and not just the errors of some moralists” (Lewis 2005, p. 318).⁵⁹

(iv) I think that the second and the third concerns are mostly problematic for the error theory. What’s more, they can be strengthened by an objection originally due to Crispin Wright (1995). We can summarize it as follows. Most error theorists are not *abolitionists*. They don’t think, that is, that we would be better off if we got rid of our moral framework. Moral concepts, unlike other concepts that embed a fundamental mistake, like “witch” or “phlogiston”, play an important role in guiding action, facilitating cooperation, and so on.⁶⁰ In light of this, it seems that moral claims can also be evaluated as correct or incorrect with respect to this subsidiary aim. But then there is a natural worry, namely: doesn’t this provide excellent reasons for thinking that we should explicate the meaning of moral sentences in terms of the satisfaction of that subsidiary aim?⁶¹

The force of this objection is not easy to assess. The error theorist could respond that

⁵⁸A conventional implicature version of the error theory would make exactly this prediction. Moreover, consider Harman’s (1996) suggestion that even if “eating fish is wrong” (W) doesn’t *mean* the same as “relative to moral framework *M*, eating fish is wrong” (W*) (where this means, roughly, that speakers wouldn’t accept that this is what they *really meant*), “for the purposes of assigning *truth-conditions*” we can understand it as (meaning) (W*). This approach can arguably be read as an instance of this attenuated error theory.

⁵⁹Joyce (2011) explores Lewis’ option and seems to sympathize with it. For a similar objection to the error theory, see also Kirchin (2010).

⁶⁰Those who think that we shouldn’t stop moralizing even if moral concepts incorporate a mistake have offered different suggestions. Some think we should retain moral discourse without altering its semantics, while others think that we should amend the meaning of moral terms, by replacing the defective meanings with some suitable successors. (See Kohler & Ridge 2013 for an overview of the various proposals, and a defence of the “replacement” strategy on behalf of the error theorist.)

⁶¹Note that this worry is more sophisticated than a blanket appeal to principles of charity and truth-maximization. This is shown by the fact that it doesn’t apply to the abolitionist error theorist.

it is precisely in order for morality to fulfill its social function that we need a language that apparently refers to objective values, or that aims to describe objective values, or that works as if there were objective values. Nonetheless, I think that Wright's observation individuates a major weakness in the metasemantics of the error theorist.

The upshot of (i)-(iv) is that even if we grant Mackie's contention that a claim to objectivity "has been incorporated in the basic, conventional, meanings of moral terms" and that "any analysis of the meanings of moral terms which omits this claim to objective, intrinsic, prescriptivity is to that extent incomplete" (p. 35), we should look in more detail whether this element can be accounted for by other semantic theories that don't predict that most moral judgments are defective.

Chapter 3

Trivial Descriptivism

3.1 Introduction

The accounts reviewed so far rely on the following assumption: an utterance u of a normative sentence φ is associated with, or determines, a *unique* proposition p , if any, which is the semantic content of that sentence in that context of utterance. This assumption leaves three roads to a variabilist who wants to stick with a propositional semantics: saying that φ expresses p_1 in c_1 and p_2 in c_2 , saying that the truth-value of p varies across assessors, or saying that p is untrue (where “untrue” is designed to cover various kinds of defectiveness, such as falsity, neither truth nor falsity, failure to express a truth-evaluable content while purporting to express one, etc.).

In this Chapter I want to articulate, explore and motivate a new hypothesis about the semantics and pragmatics of normative terms, one that drops the uniqueness assumption shared by the extant proposals. The core conjecture is that an utterance of a normative sentence is semantically associated with a *set* or *cloud* of standard-relational propositions or, more neutrally, purely descriptive propositions, which are bivalent and possess absolute truth-values.¹

¹By talking of “semantic association”, I simply want to mark a contrast with views, like the one defended by Cappelen and Lepore (2005) on which an utterance has many *speech-act contents* or *asserted contents*—“indefinitely many propositions are said, asserted, claimed, stated” (p. 4)—but only one of them is the *semantic content*. My talk of “semantic association” at this stage is neutral with respect to how exactly the *one utterance-many propositions* hypothesis is to be formally modeled (e.g. it’s not already committed to the compositional semantics itself being cloudy). All this will become clearer as we move forward, of course.

For instance, an utterance of “eating fish is wrong” is associated with the following standard-indexed propositions, where \mathbf{s}_i is a standard: {<eating fish is wrong according to \mathbf{s}_1 >, <eating fish is wrong according to \mathbf{s}_2 >, ..., <eating fish is wrong according to \mathbf{s}_n >}. Alternatively, we could say that the utterance is associated with the following descriptive propositions, where N_i is a natural property (or conjunction/disjunction of several natural properties): {<eating fish has N_1 >, <eating fish has N_2 >, ..., <eating fish has N_n >}.²

This conjecture is inspired by an analogy with a certain view about the semantics of *vague* language. On an oversimplified and unspecific version of this story, when I say “Harry is bald” there exist multiple candidate semantic values for “bald”—one for each admissible precisification of “bald”—and consequently multiple candidate “precisification propositions” associated with my utterance. A precisification is, roughly, a way of making precise the meaning of a vague expression. For instance, “is bald” can be precisified as “with fewer than x hairs” or “with fewer than y hairs”, “with fewer than z hairs”, and so on. So, an utterance of a vague sentence like “Harry is bald” is associated with the following precisification propositions: {<Harry is bald according to precisification \mathbf{p}_1 >, <Harry is bald according to precisification \mathbf{p}_2 >, ..., <Harry is bald according to precisification \mathbf{p}_n >}.³

This general story is often accompanied by two further claims: a postsemantic claim according to which a vague sentence φ is true/false iff all its precisifications are true/false, and neither true nor false otherwise; and a metasemantic claim according to which vagueness is a matter of linguistic indecision (as Lewis (1986, p. 212) famously put it “nobody has been fool enough to enforce a choice...”).³ But we don’t need any of those additional commitments to get our analogy going. All that matters is the association of a range of propositions (rather than just one) with a given utterance.

That said, there is a crucial difference between “bald” and “wrong”. Some persons

²I borrow the “cloud” metaphor from von Fintel and Gillies (2008) and (2009). They defend a *sui generis* contextualism about epistemic modals, according to which epistemic modal sentences (like “the treasure might be under the palm tree”) “put into play” a “cloud of propositions”. I will discuss their view, and compare it to my own, in section 3.5.

³This combination of claims usually goes under the name of “supervaluationism”. See Kit Fine’s seminal (1975).

are bald no matter how we precisify “bald”. No precisification is admissible that counts a man with zero cranial hairs, say Vin Diesel, as not bald, nor one that counts someone like Johnny Deep as bald. I’m misusing language if I describe the former as not bald, and the latter as bald. In other words, some vague sentences are such that they are true (or false, as the case may be) on all their admissible precisifications. In light of this, we can safely ignore their multiplicity of meanings, as it were.⁴ Normative sentences aren’t like that. Take any action you like. You may always find a moral standard which counts that action into the extension of “wrong” and one that does not, and yet such that both respect all the relevant meaning-fixing facts (use, intentions, conventions, etc.). Normative language, we could say, is *wildly* or *irresolvably vague*. In a slogan, there are no “clear cases” of application: every item x is a borderline case for a normative predicate F (or is in its “grey area”).⁵ Importantly, the analogy with vagueness is just that: an analogy. In particular, of the three features typically associated with vagueness—the presence of borderline cases, the lack of sharp boundaries and the sorites-susceptibility (see, e.g., Eklund 2007)—it is only the first that matters for my purposes.⁶

In my view, this hypothesis can be accepted by normative absolutists (invariabilists, in my earlier terminology) and by normative relativists (variabilists, in my earlier terminology) alike. What distinguishes the two is that the latter think that the competing normative standards are equally correct from a *metaphysical* point of view, and not just from the point of view of semantic/conceptual competence. It’s not just that the conventional meaning of normative words is such that they can be applied to any objects or situations whatsoever without impugning the semantic competence of the speaker. None

⁴As Lewis (1993, p. 172) puts it: “It makes no difference just what you meant, what you say is true regardless”.

⁵To my knowledge, the view that moral/normative predicates are wildly vague/indeterminate was first floated by Aqvist (1964) in a largely unknown paper titled “Vagueness and Value”. Unfortunately, Aqvist’s view is developed in the context of an old-fashioned theory of vagueness according to which borderline predications are meaningless. More recently, Schiffer (1990) and (2003) has advocated a view in the same ballpark. Schiffer’s view, though, is difficult to assess since it is framed in the context of his psychological (contrast: semantic) characterisation of vagueness/indeterminacy. The idea that normative terms are maximally vague has recently been advanced and defended in a more elaborate form by Bedke (ms). I will discuss his view below.

⁶Normative expressions can certainly exhibit the classical soritical-vagueness associated with “bald”. For an interesting approach to this problem, see Dougherty (2013). What I will discuss here, though, is largely independent from these issues.

of the worldly facts help determine the extension of the word “wrong” either. The approach I will present here is a package consisting of the (semantic) thesis that normative language is radically vague and of the (metaphysical) variabilist thesis that no normative standard is metaphysically privileged, or objectively best, or some such thing. I call this combination of views, in the specific format I will plump for, *Trivial Descriptivism*.

It is important to keep in mind, though, that the thesis that normative sentences express many propositions, some of them true and some other false, can, in principle, be paired with a postsemantics or pragmatics that fits the need of the invariabilist. (More on this in section 3.3.3). In this sense, a cloudy approach to the semantics of normative language, much like a contextualist or a relativist one, is not bound to be variabilist.

To wrap up, the general thought is that an utterance of a normative sentence is semantically associated with many propositions, some true and some false. This sketch obviously raises a number of questions, and choice-points. Two stand out as especially important. First, there is a *formal-semantic* question: how can we model the one utterance-many propositions hypothesis in the context of the semantic framework outlined in Chapter 1, and in accordance with general compositionality principles? Second, there is a question in the *pragmatics*: how does the envisaged plurality of propositions impact on the correctness/incorrectness of uses of the target sentences?

This chapter is organized as follows. In §3.2 I sketch the basic elements of a *cloudy semantics*. In §3.3 I outline the skeleton of the pragmatic framework that I’ll be adopting throughout this chapter. In §3.4 I present the full trivial-descriptivist proposal, and explain how it captures the various linguistic data concerning normative assertion, agreement and disagreement that we looked at in Chapter 2. §3.5 compares trivial descriptivism to some similar approaches, and §3.6 discusses some metasemantic issues. Finally, in §3.7 I briefly examine some alternative implementations of the cloudy hypothesis at the formal-semantic level, more precisely some cloudy treatments that keep the cloud out of the semantics.

Trivial descriptivism is a non-cognitivist view: normative sentences, when used internally and categorically, do not divide the space of possible worlds, by ruling in and out pos-

sibilities. By making a normative claim one does not describe some feature of the world, not even his attitudes/standards, or the content of his attitudes/standards. The function of these sentences is to communicate, discuss, and negotiate our attitudes/standards, but this effect is achieved indirectly, as it were. As we shall see, trivial descriptivism models this non-cognitivist insight by saying that normative utterances change the context not *via* what they *say*, but *via* what they indirectly communicate in virtue of other aspects of their meaning, broadly understood. This offers a potential advantage over views that capture the non-cognitivist idea by incorporating attitudes/standards/hyperplans/etc. in the semantic content, which is modeled, e.g., by sets of world-standards pairs. In my view, by taking standards (or the like) to be part of the semantic content, these formal representations don't fully reflect, and respect, the underlying non-cognitivist thought that standards, unlike worlds, aren't described by the corresponding judgments/assertions.

More generally, I will attempt to show how the present approach may offer some advantages over the competing accounts discussed in the previous chapter. Trivial descriptivism, unlike some versions of expressivism, lacks a burdensome commitment to an *attitude semantics*. Unlike, relativism, it makes do with the ordinary, absolute, conception of truth and has no need to “make sense of relative truth”. And unlike contextualism, it assigns a common semantic value—i.e. a set of propositions—to utterances of the same normative sentence made by different speakers. As I shall explain, this puts the view in a better position to account for a range of linguistic data concerning disagreement, attitude ascriptions/speech-reports, and retraction, which contextualists have a hard time explaining. Of course, I'm looking ahead, and I'm only highlighting the good-making features. So, let's move on, starting with the cloudy semantics.

3.2 Cloudy Semantics Explained

To some extent, a trivial descriptivist endorses the propositional semantics sketched in Chapter 1. Roughly, he affords a central role to propositions in explaining the semantic contents of sentences in context, the objects of mental attitudes and illocutionary

speech-acts, and the primary truth-bearers. He also accepts that (declarative) normative sentences, like other sentences in the indicative mood, are conventionally used to make assertions, instead of other speech-acts.

However, he departs from the standard framework in one crucial respect. He holds that the semantic content of a normative expression is *a set of intensions* (equivalently, a set of Kaplanian contents). One can think of intensions in different ways. If we adopt a coarse-grained conception, the semantic content of a normative predicate will be *a set of sets of individuals*, and that of a normative sentence *a set of sets of worlds*. The trivial descriptivist thus opts for a “cloudy semantics”, i.e. an account that builds the envisaged multiplicity of meanings into the semantics itself.⁷

Before I give a sketch of how a cloudy semantics can be formally implemented, I will clarify some more philosophical features of the present account, by dwelling a little further on the analogy with vagueness/indeterminacy in descriptive language.

3.2.1 Multiplicity and Indeterminacy

The question: “what proposition(s), if any, is expressed by a sentence that contains a vague term?” receives at least three distinct answers by proponents of linguistic accounts of vagueness.⁸

- (i) A vague sentence indeterminately expresses one of the candidate bivalent/precise propositions.⁹
- (ii) A vague sentence determinately expresses all the candidate bivalent/precise propositions.¹⁰
- (iii) A vague sentence determinately expresses a unique non-bivalent/vague proposition.¹¹

⁷As mentioned, in §3.7 I explore some pragmatic implementations of the cloudy hypothesis.

⁸The contrast is with epistemic, metaphysical, or psychological theories of vagueness.

⁹In the literature on supervaluationism, Keefe (2000) goes for this view.

¹⁰Fine’s remark that: “Vagueness is ambiguity on a grand and systematic scale” could be read as supporting (ii). There is room for reading Braun and Sider (2007) as endorsing something like this view.

¹¹Smith (2010, 2011) claims that at least some self-described supervaluationists typically adopt (iii). He uses the term “plurivaluationism” to characterize an alternative approach which could be either (i) or (ii) in our list: “On this view [on the plurivaluationist view], when I utter ‘Bob is tall’, I say *many things* at once: one claim for each acceptable model. Thus we have semantic *indeterminacy*—or equally, semantic *plurality* (2011, p. 8)”. (Smith, p. 8).

A cloudy semantics corresponds to option (ii). This proposal introduces a new, *one-to-many*, semantic expression relation between sentences and propositions. In this sense, it drops an assumption shared by (i) and (iii), namely that the expression relation is (determinately or indeterminately) of the *one-to-one* type. Meanwhile, it agrees with (iii) (and departs from (i)) in taking the expression relation to be determinate, and it agrees with (i) (and departs from (iii)) in employing only classical bivalent propositions.

I think that the difference between (ii) and (iii) is sufficiently clear. At the very least, its understanding is parasitic on one's appreciation of the distinction between bivalent and non-bivalent propositions. It will prove useful, instead, to say a little more about how (ii) differs from (i). I take it that the difference in how the content of a vague sentence is *formally represented* by the two accounts is easy to grasp. To be sure, I will adopt option (ii) largely because it arguably fits more smoothly into the orthodox propositional semantics. One question, though, is whether this formal choice brings with it some specific philosophical commitments, perhaps distinct from the ones associated with proposal (i). Let me address two potential concerns.

The first is that an account of type (ii) would imply a metasemantic analysis on which vagueness stems from meaning *overdetermination*, rather than *underdetermination* (i.e. indecision). I don't think any such consequence follows. (i), (ii) and (iii) are the "building blocks" of formal models that can be used to predict the truth-conditions of sentences or the correctness conditions of uses thereof. As such, they can be paired with different stories as to why and how the linguistic expressions in questions are endowed with such-and-such meanings.

A second worry is that if a sentence expresses true and false propositions, then that sentence is *both* true and false. Borderline vague sentences (and hence all unembedded normative sentences) would come out being true and false, a prediction that many would like to avoid. Fortunately, nothing like that is entailed by (ii). A set of propositions is *not* the same thing as a conjunctive or disjunctive proposition. It is open to an advocate of (ii), as much as to one of (i) or (iii), to define sentential/assertoric truth as truth of all/some/most of the candidate propositions (or for that matter, to not define a notion

of truth for sentences/assertions in terms of the truth of the proposition(s) they express, and to allow for a disquotational truth-predicate only).

Having clarified that, in what follows I will mostly use the terminology of multiplicity, rather than the one of indeterminacy. Besides reflecting the semantic framework adopted, the language of multiplicity, unlike that of indeterminacy, doesn't convey the thought that normative language is necessarily defective, and is compatible with the view that, to use a locution from von Fintel and Gillies (2008), the envisaged multiplicity of meanings is "by design".

It may be worth mentioning a fourth answer that, compatibly with a semantic approach to vagueness, one could give to our initial question: "what proposition is expressed by a vague sentence?" Namely:

- (iv) A vague sentence doesn't express any proposition.

One way of cashing this out without predicting that vague sentences lack content altogether, is to say that they express a proposition-*like* entity. As a general semantic theory, this is the approach favored by philosophers like Bach (1994, 2011). The idea is that the semantic contents of sentences (in context) are not full-fledged, truth-evaluable, propositions, but rather *propositional radicals* (or *propositional functions*), i.e. entities that need to be "elaborated", "completed" or "filled in" in order to express something truth-evaluable. This approach is very similar to the one pursued by a cloudy semantics: the "propositional functions" theorist says that the content of a sentence in context is a radical whose completions are all the candidate propositions in a given set; the "many propositions" theorist says that the content of a sentence in context is the set of all the candidate propositions.¹²

¹²One important difference between Bach's overall view (see, for instance his 2011) and the one pursued here is that propositional radicals are not taken to be the objects of propositional attitudes/speech-acts. On my view, on the other hand, sets of propositions play both the semantic content role and the asserted content role.

3.2.2 A Toy Cloudy Semantics

A bit of stage-setting and terminology will help. According to a fairly canonical approach,¹³ a semantics for a natural language consists in a definition of an interpretation function $\llbracket \cdot \rrbracket$ which assigns an extension to expressions of the language relative to a model M , a variable assignment function g , a context of utterance c , and an index or circumstance of evaluation i . So, using “ α ” as a schematic letter for sentential and subsentential expressions, “ $\llbracket \alpha \rrbracket^{M,g,c,i}$ ” describes the extension of α relative to a model M , an assignment function g , a context c and an index i .¹⁴ For simplicity, in what follows, I will often omit both M and g . The function will thus simply be: “ $\llbracket \alpha \rrbracket^{c,i}$ ”.

The double bracket function $\llbracket \cdot \rrbracket$ is an *extension* function. However, one can define an *intension* function in terms of it. For example, assuming that the index i only contains a parameter w that ranges over worlds, so that we have $\llbracket \alpha \rrbracket^{c,w}$, we can calculate for any context c , the function from a given world w to the extension of α in w . I will use $|\alpha|^c$ to represent the intension of α as interpreted in a context c .

We are now in a position to articulate the cloudy proposal. The most straightforward way of generating cloudy semantic contents is perhaps the following. First, we enrich indices so that they include not only worlds, but also normative standards. So, indices are pairs $\langle w, s \rangle$ of worlds and normative standards, and the denotation function is thus $\llbracket \alpha \rrbracket^{c, \langle w, s \rangle}$ where $w \in W$ and $s \in S$, and $W =$ the set of possible worlds, and $S =$ the set of possible moral standards. Consider our usual example:

- (1) Eating fish is wrong

The extension of (1) can be represented as follows:

- (2) $\llbracket \text{Eating fish is wrong} \rrbracket^{c, \langle w, s \rangle} = 1$ iff eating fish is wrong according to s in w .

¹³See, among others, Chierchia and McConnell-Ginet (2000).

¹⁴A model M consists of a domain of discourse D and of an interpretation function I , which maps each individual constant to an object in D , and each predicate constant to a set of ordered n -tuples in D . On the basis of the interpretation function I and of a set of compositional rules, a generalized interpretation function $\llbracket \cdot \rrbracket^M$ is built up recursively that assigns to every expression (not just to simple ones) its extension $\llbracket \alpha \rrbracket^M$, i.e. the extension of α in M . Thus $\llbracket \cdot \rrbracket^M$ extends the function I of M . The assignment function g is needed in order to take care of variables (e.g. pronouns).

Thus, the standard definition of the denotation function $\llbracket \cdot \rrbracket$ is modified in the same way as proposed by MacFarlane (2014). Extensions are assigned relative to contexts and $\langle w, s \rangle$ pairs. The next step, and here is where the view departs from the relativist framework, consists in defining a notion of the content or intension of a sentence. This can be done as follows:

$$(3) |\alpha|^c = \{x \mid x \in W \text{ and } \exists s \in S: \forall w \in W, w \in x \text{ iff } \llbracket \alpha \rrbracket^{c, \langle w, s \rangle} = 1\}.$$

In other words, the intension of an expression α in a context c is the set of sets x that meet the following condition: x is a set of worlds, and for some standard s , x contains all worlds such that $\llbracket \alpha \rrbracket^{c, \langle w, s \rangle} = 1$. By using (3) we can generate contents that are sets of classical possible world propositions. For example, the intension of (1) in a context c can be represent as follows:

$$(4) |\text{Eating fish is wrong}|^c = \text{the set of sets of worlds } x \text{ such that for some standard } s, \text{ and for all } w \in x, \text{ eating fish is wrong according to } s \text{ in } w.^{15}$$

If we limit ourselves to the logical connectives, the extensions of more complex sentences can be obtained as follows:

- $\llbracket \neg \varphi \rrbracket^{c, \langle w, s \rangle} = 1$ iff $\llbracket \varphi \rrbracket^{c, \langle w, s \rangle} = 0$.
- $\llbracket \varphi \wedge \psi \rrbracket^{c, \langle w, s \rangle} = 1$ iff $\llbracket \varphi \rrbracket^{c, \langle w, s \rangle} = \llbracket \psi \rrbracket^{c, \langle w, s \rangle} = 1$.
- $\llbracket \varphi \vee \psi \rrbracket^{c, \langle w, s \rangle} = 1$ iff $\llbracket \varphi \rrbracket^{c, \langle w, s \rangle} = 1$ or $\llbracket \psi \rrbracket^{c, \langle w, s \rangle} = 1$.

Obviously, we should now give a definition of content for these expressions too. To simplify, I will use “ p_1 ”, “ p_2 ”, ..., “ p_n ”, and “ q_1 ”, “ q_2 ”, ... “ q_n ” to refer the single possible

¹⁵There may be other ways of generating contents that are sets of propositions. In particular, there may be strategies that don't require having systems of standards as extra parameters in the index. A compositional semantics that uses structured propositions, for instance, might achieve this result. But compositional semantics that employ structured propositions face many problems, and it's unclear whether there is any consistent theory as yet. A different option could be to say that the semantic content of a normative sentence not a set of set of worlds, but a set of sets of contexts. For instance, in the case of “eating fish is wrong” this would be the set which contains all sets of contexts such that eating fish is wrong according to the standard of that context. Again, this might be an option, but it requires many revisions to the existing framework where propositions are treated as sets of indices, rather than sets of contexts. In any case, I leave this topic for future research.

world propositions that are members of the set that is the semantic content of φ and ψ respectively. So, we will have something like this:

- $|\neg\varphi|^c = \{W \setminus p_1, W \setminus p_2, \dots, W \setminus p_n\}$
- $|\varphi \wedge \psi|^c = \{p_1 \cap q_1, p_2 \cap q_2, \dots, p_n \cap q_n\}$
- $|\varphi \vee \psi|^c = \{p_1 \cup q_1, p_2 \cup q_2, \dots, p_n \cup q_n\}$

To exemplify, consider the complex sentence: “Eating fish is wrong and torturing babies for fun is wrong”. Its intension can be represented as follows:

- (5) $|Eating\ fish\ is\ wrong\ and\ torturing\ babies\ for\ fun\ is\ wrong|^c =$ the set of sets of worlds x such that for some standard s , and for all $w \in x$, eating fish is wrong according to s in w and torturing babies for fun is wrong according to s in w .

If you are wondering why I’m recovering the intensions of conjunction and disjunction by intersecting/unioning the first member of the intension of $|\varphi|^c$ with the first member of the intension of $|\psi|^c$, the second with the second, etc. instead of having their cartesian product (i.e. $p_1 \cap q_1, p_1 \cap q_2$, etc.) my reply is this. For simplicity, I’m idealizing over inconsistencies and indecisions in my characterization of a standard. So, there won’t be a standard that assigns distinct extensions (consistent or not, as the case may be) to “wrong” when the predicate is applied to different actions (e.g. eating fish and torturing babies). Clearly, no real-world agent is maximally decided and consistent, so a full story will have to represent indecision and inconsistency in subjects’ normative standards (plausibly, by saying that subjects’ mental states are best represented by a set of standards, rather than by a single standard). But we can ignore this for our illustrative purposes.

Like I said, the framework I provided is meant to be a “toy” cloudy semantics. Among other things, I’ve only considered an adjective (i.e. “wrong”). I think the same strategy can be generalized to terms belonging to different syntactic categories (modals, sentential operators, nouns, etc.), but a comprehensive account is beyond the scope of the current investigation.

I will conclude this sketch by saying something about how normative sentences embed under propositional attitude verbs, like “believes”. Consider a simple belief ascription:

(6) Mia believes that eating fish is wrong.

Intuitively, an utterance of (6) communicates that Mia endorses a standard that forbids eating fish.¹⁶ How can our cloudy semantics predict this? I suggest the following:

(6) $\llbracket \text{Mia believes that eating fish is wrong} \rrbracket^{c,i} = 1$ iff Mia is committed to a standard s_i such that the corresponding proposition in the content expressed by the complement clause, i.e. $\langle \text{eating fish is wrong according to } s_i \rangle$, is true in all worlds compatible with what Mia believes.

Importantly, a cloudy semantics will have to assign sets of intensions to linguistic expressions in general, rather than to normative language only. The reason is that, plausibly, we want a unified semantics for the whole language. Normative and descriptive language can mingle in a single sentence, after all, and we don’t want to postulate ambiguities to account for that. A cloudy semantics will thus assign sets of intensions (in context) to expressions belonging to all areas of discourse; if an expression has a unique intension, its intension will be the singleton set $\{p\}$.

3.2.3 Some Worries and Replies

One might think that the current proposal is in violation of theoretical *conservatism*, for it introduces new semantic values (or a new character-function) based only on the behaviour of a fragment of language. I don’t think this is true. For one thing, the account is conservative in the respect that many people care about: it construes contents merely from classical propositions. Additionally, though this is slightly more controversial, it is widely acknowledged that most natural language sentences are vague, and a cloudy semantics may be well-suited to handle that. Finally, cloudy semantics have already

¹⁶I will consider some complications in §3.2.3 and §3.4.1.

been proposed to deal with other linguistic phenomena.¹⁷ This suggests that there may be independent motivations for pursuing a project along those lines, by exploring the idea that sets of propositions may be adequate tools for representing the meaning of sentences.

One might also worry that the current proposal makes for a less *unified* semantics for modals and terms of other categories, if compared to the ordinary contextualist account.¹⁸ Again, I don't think this is true. The conventional meaning of a sentence involving terms like “ought”, “right”, “bad”, etc. φ is a function from a context c to a set of propositions \mathcal{P} : $\{p_1, p_2, \dots, p_n\}$. The context distinguishes between normative and non-normative readings, as well as between the various normative interpretations (e.g. instrumental, legal, moral, all-things-considered, etc.) When φ gets, say, a moral reading, the set contains many propositions; when given, say, an instrumental reading it contains only one proposition.

A related worry is that the present approach might be less flexible than contextualism in accommodating what we called “external uses” of moral or other normative terms. Take the following scenario. Jim and Mia are discussing ancient Greek history. Jim says: “Pericles (morally) ought not to have ostracized Cimon” (O). In saying that, Jim need not be endorsing any of Pericle’s moral values. Rather he may simply be assuming his perspective and making a judgment from that perspective. The worry is that this implies that the kind of content that Jim expresses by using (O) in that context c is different from the kind of content that would be expressed by (O) in a context c' in which (O) receives an internal interpretation. This is a genuine worry. However, there are different things one could say in reply.

The first would be to deny that external readings are available, and to posit a

¹⁷Sets of propositions are often used to model the meaning of *questions* (cf. Hamblin 1973). In the case of declarative sentences, some have proposed to assign sets of meanings to *ambiguous* expressions (cf. Poesio 1996). In the philosophical literature, a number of authors have suggested dropping the assumption that a sentence-in-context semantically expresses at most one proposition. For instance, Bach (1999) and Neale (1999) claim that an utterance of “Mia is poor, but honest” semantically express two propositions—(a) that Mia is poor and honest; (b) that there is a contrast between being poor and being honest—and extend the multiple propositions framework to other phenomena too. King (2015) points out that in many cases speakers’ intentions determine “a range of potential semantic values”, rather than a single value, for a context-sensitive expression. Moreover, in the literature on probabilistic beliefs it is common to represent mental states by sets of probability functions (cf. Rothschild 2012). And at least some versions of supervaluationism about vagueness can be understood as rejecting the uniqueness assumption as well (although the issue is somewhat tricky, since supervaluationists don’t say very much about the propositional contents of vague sentences. See the discussion in the previous paragraph).

¹⁸See the discussion in Chapter 2.

constrain on the interpretation of normative sentences that the speaker's standard is always among the relevant standards. In this sense, a moral term would be similar to an indexical like "we". We cannot felicitously say things like: "we went to the grocery, but I didn't go".¹⁹ This maneuver strikes me as weak. Even if we grant that "Pericles (morally) ought not to have ostracized Cimon, but I don't disapprove of his action" (O*) is anomalous (something which I take to be controversial) this solution is not easily generalizable to other varieties of normative claims.²⁰

A more promising route could be to accept that "morally ought" has a different content when used externally from the one it has when used internally.²¹ In internal uses it expresses a set of propositions containing many propositions, one for each possible moral standard. In external uses it expresses a set which only contains one proposition, the one corresponding to the salient external standard. This splitting in the semantics can be motivated by noting that external readings of utterances like (O) are hard to get, as proved by the fact that claims like (O*) are somewhat stilted, even if not incoherent. A third option is to say that (O) expresses the same set of propositions in both internal and external uses. The difference lies in the conditions under which it is correct for someone to believe or assert such a set.²²

So, it seems to me that a trivial descriptivist could choose between the second and the third option to handle external uses of normative words. To further mitigate lingering worries, it may also be important to note that contextualists too will have to tell some story to allow for external uses, and to distinguish them from internal ones. So even if, compared to our second option, a contextualist account is, in effect, more unified—"wrong" always means: "forbidden by the salient standard"—the contextualist too should

¹⁹Weatherson and Egan (2011) argue that there is a "speaker inclusion constraint" on the interpretation of epistemic modal claims, to the effect that the speaker's knowledge is always among the relevant epistemic standards.

²⁰For instance, to use a case invoked in the literature on predicates of personal taste, the shop assistant who sells food for animals may call "tasty" a new brand without thereby giving voice to her own culinary preferences.

²¹MacFarlane (2014, pp. 155-156) proposes an analogous strategy to deal with external uses of predicates of personal taste.

²²Among proponents of a relativist semantics for predicates of personal taste, this is the option favored by Lasersohn (2005). Lasersohn's approach requires a new semantics for propositional attitudes on which "believe" is treated as a "3-place relation between an individual, a context, and a sentence content" (Lasersohn 2005, p. 676).

provide an explanation and representation of the fact that the salient standard may be, and in fact typically is, internal (i.e. accepted in the conversational context) or external (i.e. not necessarily accepted by the conversationalists).

These remarks would need a fuller elaboration, of course. However, my primary interest here is to offer an account of *internal* and *categorical* uses of normative language. The point of the foregoing considerations was to show how the present proposal may fit into some standard linguistic frameworks (for modals and other terms) and to take a preliminary step in the direction of a broader uniform semantics for normative language. With this in mind, I will now turn to my chief task.

3.3 Pragmatics: A General Framework

Ultimately, the semantic facts should allow us to make the right predictions about the correctness and incorrectness of language users' speech-acts. This is their primary job, at least if we stick to a certain conception of natural language semantics (see Chapter 1, §1.5). One of the motivations behind a cloudy account was that it promises some theoretical advantages over its competitors; more cautiously, that it offers a new, empirically adequate, model for explaining the relevant linguistic data. I'm now going to explain how the "cloudy semantic facts" bear on explaining the way in which the target sentences are conventionally used in conversation, and relatedly on predicting the various norms governing the production, acceptance, rejection, and retraction of normative claims in the dynamics of typical conversations.

To this end, I will articulate a *dynamic pragmatic* framework which will underlie my overall approach. More specifically: §3.3.1 introduces a "conversational scoreboard" model of context; §3.3.2 spells out the idea that assertions, like other speech-acts, are proposals to change the context; §3.3.3 takes up some outstanding issues at the semantics-pragmatics interface, in particular the question of truth *simpliciter* or monadic truth.

3.3.1 Contexts as Scoreboards

The first point concerns how to think about context. Stalnaker (2014) distinguishes between three notions of context: an *informal* notion, according to which a context is the concrete situation in which a conversation takes place and two *formal* notions, i.e. two ways of modelling or representing that concrete context, by means of ordered n -tuples of parameters. The first, call it c , is the Kaplanian context, a quadruple consisting of the agent, the time, the location and the world of the utterance. The second, call it cs , is the notion of *conversational score* or *common ground*, introduced by Lewis (1979) and Stalnaker (1978), and developed in theories of dynamic semantics.²³

There are two basic and related ideas behind this notion of context. The first is that conversations normally take place against a background of mutually shared commitments—in Lewis' words: "Sentences depend for their truth-value, or for their acceptability in other respects, on the components of conversational score at the stage of conversation when they are uttered" (1979, p. 345). The second is that utterances, besides being context-sensitive, can have a context-changing effect—again, in Lewis' words, the conversational score "tend[s] to evolve in such a way as is required in order to make whatever occurs count as correct play" (p. 347).

The conversational score includes, at least, a body of information taken for granted, or presupposed, by the discourse participants. But since conversations are not only aimed at communicating information or ruling out worlds, and since cs can be changed by speech-acts other than assertions (e.g. commands and questions), we may enrich the representation of the score by adding further parameters in addition to the information-tracking one. For example, Lewis (1979) takes "standards of precisions" and "spheres of permissibility" to be a component of cs and Roberts (2002) introduces the idea of a "set of questions under discussion". More importantly for our purposes, we can imagine a parameter that keeps track of the conversationalists' mutually agreed-upon practical

²³In the framework of dynamic semantics, the semantic values of sentences are *context-change potentials* rather than truth-conditions. It is common in the literature to distinguish between a "dynamic pragmatics" and a "dynamic semantics". Stalnaker (2014) describes the difference as follows: "The former says that content is determined as a function of context, and then a general force rule explains how the content changes the context. The latter streamlines the story, eliminating the middleman by building force into the semantic values of sentences". See also K. Lewis (2014) for discussion.

commitments or systems of standards.²⁴

Obviously, c (i.e. the Kaplanian context) and cs complement each other. For example, we can think of cs as somehow part of, and determined by, c .²⁵ I will sometimes use $cs(c)$ to refer to the conversational score of the context c .

3.3.2 Assertions as Proposals

A pragmatic theory, conceived of as a theory of speech-acts, has two main goals. The first is to investigate possible formal models of the context or conversational scoreboard, and to describe the ways in which different speech-acts change or update the context. This bit of the enterprise is often referred to as “formal pragmatics”. The second is to articulate the norms that govern the production of speech-acts, and thus the conversational dynamics. In the case of assertion, this typically involves an inquiry into the “norms of assertion”. More broadly, it has to do with the “felicity conditions” of various speech-acts: under what conditions is someone entitled to make an assertion, issue a command, ask a question, make a promise, and so on? In this section, I will sketch a view that provides a distinctive answer to both questions.

The dynamic approach to context briefly outlined in the previous section has a companion theory of assertion. According to a *minimal* version of this theory, assertions, like other speech-acts, are proposals to change the value of some component of the conversational score. What is distinctive of assertions *vis-à-vis* other conversational moves, such as commands and questions, is that they update the context by intersecting the conversational score—which we can simply represent as a set of propositions—with the content of the assertion. To be a little more precise, if cs is the score and $|\varphi|^c$ is the content of the assertion, we have: $cs \cap |\varphi|^c$.²⁶

²⁴I prefer Lewis’ terminology of “a conversational score” to Stalnaker’s “common ground”, because the latter is often used in a narrow sense, to talk about the *information* presupposed.

²⁵See Stalnaker (2014), pp. 64-65. Alternatively, one could think of cs as a set of c -contexts (*ibid*).

²⁶The reason why I’m calling this approach “minimal” is that, unlike in the original Stalnakerian framework, assertions can update the context in ways that involve something else, or something more, than narrowing down the set of live possibilities. For instance, it is compatible with this minimal account that assertions change the context by *expanding* rather than shrinking the set of open possibilities (as epistemic modal claims like “might φ ” do on some relativist views) and by ruling out *more fine-grained entities* than just worlds, like sets of $\langle \text{world-norm} \rangle$ pairs (as normative claims do on some relativist views).

For instance, to take the simplest case, if φ is used to make an ordinary factual claim, an assertion of $|\varphi|^c$ updates the score by ruling out the worlds that are incompatible with its content. Equivalently, an assertoric utterance of φ in a context c is a proposal that the conversational score entails the intension of φ in c .

Importantly, assertions can have *pragmatic* effects on the conversational score, in addition to semantic effects. For example, if an assertoric speech-act conversationally implicates or presupposes that p , and the assertion is accepted by the interlocutors, then p is added to the score, and becomes common ground.

The idea that assertions are proposals to change the context, rather than something else, is important for our purposes because proposals can be subject to relatively cheap and local norms. The felicity conditions of a proposal, like that of an imperative (“Shut the door!”) or a performative (“I promise you to do x ”), need not be tied to the truth-related requirement that is normally assumed for declarative sentences. Plausibly, one is licensed to make a certain proposal, if one reasonably believes, or perhaps knows, that it is a good proposal. Of course, what counts as a good proposal turns on the domain and the goals of the conversation—it is easy to see why, if we are engaged in an inquiry about the actual world, a good proposal will be subject to truth-related norms. But when our goal is, say, to coordinate on a certain set of standards, or to fix the boundaries of a certain concept, a good proposal may simply be subject to the norm that one’s proposal is in line with one’s preferences, ends, and so on. This will become important as we move on.

With this general framework in place, we have two main tasks. The first is to adjust the conversational model outlined in a way that fits the cloudy semantics sketched above. Relatedly, we should explain how, within this new setting, factual assertions change the context. The second is to explain how normative assertions change the context. The rest of this section is dedicated to the first task.

One *caveat* first: when I talk of “factual assertions” I refer to assertions that are used to describe or represent the world (perhaps with respect to some contextual parameter); when I talk of “normative assertions” I refer to assertions that are used to propose how

the world ought to be (or how a certain contextual parameter ought to be). This means that not all assertions of (unembedded) descriptive sentences are bound to be factual (I can use a sentence containing a vague term to propose a certain precisification, rather than to describe a way the world is)²⁷; conversely, not all assertions of (unembedded) sentences containing terms like “good”, “ought”, “right”, etc. are bound to be normative (assertions where “good” and “ought” are used instrumentally may well be purely factual; similarly, “external uses” as discussed in Chapter 2 and above, are clearly not normative).

So let us turn to our first goal. We achieve this goal by modifying the basic update rule for assertion. Recall that a cloudy semantics applies across the board, i.e. to both descriptive and normative language. Now, many other types of sentences, in addition to normative ones, express sets of propositions containing many propositions. One example is that of vague or imprecise sentences—like “Harry is bald” or “Harry arrived at 3pm”, where in fact he arrived at 3.01 pm. I suggest that the score is changed by adding the *disjunction* of the propositions in the set. There are two main considerations behind this choice. The first is that the propositions may be mutually incompatible. To use the examples just given, in the case of “Harry is bald” the various propositions will be compatible insofar as they reflect the various precisifications for “bald”. However, in the case of “Harry arrived at 3pm” we can imagine a range of admissible interpretations which are mutually incompatible, for instance: <the train arrived at 2.58pm>, <the train arrived at 2.59pm>, <the train arrived at 3.00pm>, <the train arrived at 3.01pm>. In this case, adding the conjunction would make the conversational score inconsistent. The second observation is that, no matter whether they are compatible or not, I think that the relevant multiplicity/indeterminacy/indecision is better modeled by saying that what is added to the conversational score is the disjunction, rather than the conjunction, of the propositions in the set. Let’s think through this by means of a toy example.

Consider the usual vague sentence: “Harry is bald”. For illustrative purposes, imagine that “bald” has three admissible precisifications—*with fewer than 100.000 hairs*, *with fewer than 90.000 hairs*, *with fewer than 80.000 hairs*—so that the set of precise proposi-

²⁷I will expand on this below.

tions expressed by our sentence is the following: {<Harry has fewer than 100.000 hairs>, <Harry has fewer than 90.000 hairs>, <Harry has fewer than 80.000 hairs>}.²⁸ An assertoric utterance of “Harry is bald” communicates *that either Harry has fewer than 100.000 hairs or Harry has fewer than 90.000 hairs or Harry has fewer than 80.000 hairs*. Equivalently, an assertoric utterance of “Harry is bald” is a proposal to rule out those possible worlds in which none of the precise propositions in the cloud is true. So, we have the following update rule:

- (7) An assertion of $|\varphi|^c$ is a proposal to add to the conversational score at c the disjunction of the members of $|\varphi|^c$.

It is important to dwell a bit on some features of the present account. When I say that what is added to the score, or what is communicated, is the disjunction of the propositions in the set I’m not saying that the semantic content or the assertoric content of the sentence is a single disjunctive proposition. In the framework I’ve opted for, the semantic content of the sentence is a set of propositions, and the assertoric content coincides with the semantic content. A *multiple* assertion, i.e. an assertion of a set of propositions, is not the same thing as a *disjunctive* (or *conjunctive*) assertion. One might wonder whether the point is not merely terminological. Call “assertoric content*” the content that is added to the conversational score. The assertoric content* of an utterance of “Harry is bald” is the disjunction of the various propositions in the set. This may well be a matter of terminology. Still, I think we might need both notions (i.e. the non-starred one and the starred one), and for the following two reasons.

First, intuitively, someone who believes/asserts the content expressed by “Harry is bald” doesn’t believe/assert the proposition <Harry is bald on some precisifications>, nor for that matter the proposition <Harry is bald on all precisifications>.²⁹ The disjunctive or conjunctive propositions are perhaps best understood as “proxies”, in the absence of a

²⁸This is an oversimplification in various respects. For instance, the supervenience base is plausibly not exhausted by the *number* of hairs on one’s head—facts about the *distribution* of hairs are also relevant. Second, and more importantly, I’m pretending we can list the various precise propositions, something which we cannot really do because of the phenomenon of higher-order vagueness.

²⁹I mention the conjunctive option for those who think that the problem might lie in the disjunctive treatment.

unique precise proposition. Second, if we say that the assertoric content is the disjunction of the propositions in the set we are forced to say that two contradictory borderline vague sentences are both true, something which raises additional worries. (I will discuss this second point more extensively in the next section.)

With this in mind, the next step is to articulate some principles of score-change for normative assertions, and then some rules for assertion that allow us to predict the various linguistic patterns described in the previous chapter. Before I do that, I will briefly elaborate on the role that the notions of truth and falsity play in my account.

3.3.3 Monadic Truth

On a cloudy semantics, a sentence may express many propositions. Propositions are the primary truth-bearers, and each proposition in the set is assigned a unique classical truth-value (i.e. true or false). However, in ordinary talk we don't assign truth-values distributively to each of the propositions in the set. Rather, we ascribe truth and falsity to the contents of speech and thought (and perhaps derivatively to other truth-bearers), which in a cloudy semantics are *sets/clouds* of propositions. How can we accommodate everyday ascriptions of truth and falsity in this new framework?

According to the orthodox truth-conditional framework, sentences-in-context, utterances, assertions, beliefs, judgments, etc. inherit their truth-conditions from the propositions they express. It arguably follows from this picture that if a sentence expresses a set of propositions, rather than a single proposition, the sentence inherits the corresponding set of truth-conditions. Thus, unlike in the standard framework, there seems to be no straightforward derivation of sentential or assertoric truth-conditions from semantic content.

However, it is crucial to note that insofar as propositions have their truth-values relative to worlds, or to n -tuples of parameters (e.g. times, standards, states of information, etc.), the question: "when is the content of the belief/assertion true *simpliciter*?" equally arises in the classical non-cloudy framework. In the case of worlds, for instance, it is common to identify ordinary propositional truth, call it truth_O , with truth_P at the *ac-*

tual world (where “ truth_P ” stands for the semantic notion of propositional truth). When subjective factors are involved, as in the case of vagueness or normativity, the issue is trickier. Which language or normative standard is the privileged one? *Prima facie*, there is no suitable candidate to play the role of the actual world (something like the actual language or the actual standard). There are two main options (let’s focus on standards):

- (i) Define a suitable monadic truth-predicate for propositions (either in terms of truth at some privileged standard, or in terms of truth at some/most/all/what-have-you, standards).
- (ii) Do not define a monadic truth-predicate for propositions. Define instead a notion of *correctness* for beliefs/assertions (the options can match exactly the ones mentioned in (i)) Ascriptions of truth and falsity are accommodated by exploiting the disquotational properties of the truth-predicate (e.g. if it is correct for A to assert $\langle p \rangle$, it is thereby correct for A to assert $\langle p \text{ is true} \rangle$).

There is a parallel choice to make for a cloudy semantics. Namely: should we define a monadic truth-predicate for our new contents, i.e. clouds of propositions? Or should we simply articulate some pragmatic norms to predict when it is correct/incorrect for someone to assert/believe such a cloud? I prefer to follow the second route. Namely, I will not define a truth notion for contents in terms of propositional truth. I don’t think that the first route is hopeless. However, for one thing it raises additional challenges that are instead escaped by an account of type (ii). For example, it involves positing truth-value gaps (if we go for the “all” options) or gluts (if we go for the “some” options). What’s more, I think that no such definition is needed to achieve the explanatory goals I’ve set for myself.

In §3.1 I said that the idea that a normative sentence expresses a set containing true and false propositions can in principle be adopted by an invariabilist, i.e. by someone who thinks that there is a uniquely adequate normative standard. We are now in a position to see how this is possible. Plausibly, an invariabilist will say that the content of a normative claim is true iff the proposition in the set that corresponds to the objectively

correct normative standard is true.³⁰ Likewise, an error theorist could adopt the same definition of normative truth and deny the existence of an objectively correct moral standard (equivalently, he could say that the content is true iff all the propositions in the set are true).

3.4 Trivial Descriptivism and the Explanatory Challenges

We now have the tools to explain how the trivial descriptivist can meet the explanatory desiderata. In particular, if you recall the data discussed in the previous chapter, the task is to explain why we are entitled to make assertoric speech-acts with normative sentences (call this **Assertion**), why asserters of contradictory normative sentences disagree (call this **Disagreement**) and why we may reasonably retract an earlier normative assertion in the light of a change in our views (call this **Retraction**).

The model I will offer has the following two features:

- (i) *Trivial descriptive content*: The semantic content of a normative sentence in context is a set of true and false propositions. For this reason, the descriptive/representational content of a normative sentence is somehow trivial: it doesn't rule out any world.
- (ii) *Pragmatic score-change*: Assertions of normative sentences have a pragmatic, rather than semantic, effect on the conversational score. The score is changed indirectly, by exploiting some pragmatic principles.

The view I will sketch represents one way of articulating point (ii). There may be other, better or worse, ways of doing that, while sticking to (i), and I will mention some of them in passing.

³⁰Compare MacFarlane's (2014, p. 67) example of an absolutist postsemantics, where the truth of aesthetic sentences is relativized to God's aesthetic standard.

3.4.1 Assertion

Start with a simple model (we will revise it in due time). Think of the conversational score as having only an information-tracking parameter. Contextual information gets updated by any unopposed assertion. In general, when a content (i.e. a set of propositions) is asserted, the contextual information is updated by adding the disjunction of the propositions in the set to the common ground. This works well with vague or imprecise assertions, insofar as they are used to communicate information. In the case of assertions of normative sentences, though, adding the disjunction is not going to add information (i.e. it is not going to rule out any possibilities). The reason is that there is no world where *none* of the propositions in the cloud is true. (Recall that the trivial descriptivist holds that for any normative ascription there is always a standard that verifies it and one that falsifies it).

However, there is an expectation that people only assert such a set when the subset containing those propositions corresponding to the range of systems of standards that they are committed to contains only propositions they take to be true. Thus, when such a set of propositions is asserted, the contextual information is not changed by updating in the normal way. It is instead changed through the extra expectation: if the sentence uttered is “eating fish is wrong” the information that the asserter is committed to a system of standards that prohibits eating fish is added to the contextual information.

We need to explain the source of the relevant presumption. Namely, where does the expectation come from that people who assert a cloud \mathcal{P} of propositions: $\{ \langle \text{eating fish is wrong according to } \mathbf{s}_1 \rangle, \langle \text{eating fish is wrong according to } \mathbf{s}_2 \rangle, \dots, \langle \text{eating fish is wrong according to } \mathbf{s}_n \rangle \}$ are expected to be committed to a system of standards \mathbf{s}_i such that $\langle \text{eating fish is wrong according to } \mathbf{s}_i \rangle$ is true at the world as they believe it to be?

There are various options (which can plausibly merge). It might be an expectation on the corresponding belief together with a presumption of *sincerity*: namely, that the speaker asserts a normative content only if those members of the content that are about the range of systems she is committed to are true given what the speaker believes. (On this line, flat out asserting that eating fish is wrong, while having no attitude whatsoever

against eating fish, would be similar to lying.).

Alternatively, or additionally, it might be clear from the context that the cloud asserted contains some propositions that are practically interesting for the speaker/thinker, namely the ones where s_i is a standard accepted by the thinker/speaker. Or yet, it might be that given the general goal of conversations about normative matters, namely to influence actions and decisions, an assertion of a normative content would lack in *relevance* if the speaker didn't accept a system of standards that verifies the corresponding propositions in the cloud.³¹

So, given that the normal update rule is unproductive, the cooperative principle requires that there be a further aim. In particular, the expectation that the maxim of quality and/or relation is obeyed would give raise to something akin to a Gricean *generalized conversational implicature* that the asserter is committed to the relevant system of standards. Let's flesh this out a little further.³²

Grice (1975/1989) famously distinguished between *particularized conversational implicatures*, *generalized conversational implicatures*, and *conventional implicatures*. According to Grice, what distinguishes conversational from conventional implicatures is the fact the former, unlike the latter, are *cancellable* and *non-detachable*.³³ There are at least four points that speak in favor of a conversational (contrast: conventional) analysis of the implicatures of our concern. Both Finlay (2005) and Strandberg (2012) convincingly point them out. The first is that the implicature in question seems highly non-detachable: it's hard to find a different coextensive term that doesn't carry the implicature that, say, "wrong" carries. The second is that there seem to be contexts in which the implicature can be felicitously cancelled. (This can happen with cold and reverse uses of normative terms, as I will discuss below).³⁴ The third is that conventional implicatures typically *project*, i.e. are inherited by complex constructions in which the original sentence is embedded,

³¹See Strandberg (2012a) for an explanation of the "practicality" of normative claims along these lines.

³²Implicature accounts have been offered by quite a few metaethicists. See, for instance, Copp (1995, 2001), Finlay (2005, 2014), and Strandberg (2012).

³³An implicature is non-detachable if "it is not possible to find another way of saying the same thing, which simply lacks the implicature in question" (Grice 1975, p. 39).

³⁴In fact, I think that a conventional implicature analysis could be able to accommodate the cancellability data by complicating the content of the implicature. More on this below.

but this doesn't seem to happen with conative attitudes ("If eating fish is wrong, then cooking fish is wrong" need not convey any attitude.) The fourth is that, in general, the very existence of conventional implicatures is taken to be controversial (see Bach 1999).

A *generalized* implicature differs from a *particularized* one in that, despite not being encoded in the conventional meaning of the sentence/word, it is highly standardized, so much so that one may call it a "convention of use".³⁵ Unlike particularized ones, generalized implicatures do not need any contextual features to arise, and are generated by default, unless there are some special defeating circumstances. I think this model nicely captures the fact that the connection between normative utterances and expression of conative attitudes is tight, "meaning-like" (Strandberg 2012), but at the same time defeasible.

I hesitate in describing this as a "Gricean implicature", for there is at least one respect in which the present mechanism differs from the one described by Grice. On the Gricean story, when A implicates that p , A typically believes that p . The supposition that A believes that p and that A is being cooperative are precisely required to derive the implicature (cf. 1975, p. 31). But this doesn't fit our case. For one thing, the speaker need not have the meta-belief that the standard he accepts forbids eating fish. For another, the presumption that he does is not required to make the relevant inference. The problem, I think, is that Grice thought that if something is trivially generated by the presumption that the speaker is cooperative, then it's not an implicature. So, for instance, he thought that the belief that p is not implicated by an assertion that p , because it *just is* the assumption that the speaker observes the quality maxim.³⁶ But we can distinguish between trivial and non-trivial implicatures (where Grice would reserve the term "implicature" to the latter) and use the term "implicature" in this extended sense to describe the present mechanism.³⁷

³⁵As Searle (1975), for instance, does. See also Bach and Harnish's notion of "standardized indirect speech acts" in their (1979).

³⁶See Grice (1989), p. 82: "On my account, it will not be true that when I say that p , I conversationally implicate that I believe that p ; for to suppose that I believe that p (or rather think of myself as believing that p) is just to suppose that I am observing the first maxim of Quality on this occasion."

³⁷Copp (2001) and Finlay (2014, pp. 119-120) similarly think that the mechanism is different from the Gricean one (Copp introduces the label "simplicature" to distinguish the pragmatic mechanism he invokes from the Gricean one). However, their motivation—a speaker need not *mean/intend* to convey

I should perhaps point out that the idea that attitudes are communicated via implicatures or quasi-implicatures is not without problems. A common complaint is that an implicature view is at best incomplete, insofar as it covers the case of normative speech, but not that of normative *thought*. As far as I can see, this worry doesn't touch the account I've proposed. For above (cf. §3.2.2) I have indeed committed myself to the view that to believe the content expressed by a normative sentence is to be in some non-cognitive mental state.³⁸ I will detail that proposal in a moment. A further issue is whether the model of implicatures offers the best model to capture the phenomenon at stake. Why not opt instead for a view on which attitudes are expressed in virtue of being part of a *Fregean sense* of the term, or perhaps of their *Kaplanian character* (as suggested in Dreier 1990), of normative words? I think these may be attractive routes to pursue, especially the latter, but I leave this topic for future research.³⁹

In light of the foregoing considerations, we can articulate the following norm for belief/assertion:

- **Belief/Assertion:** An agent A is entitled to believe/assert the cloud $|\varphi|^c$ only if he accepts a standard s_i such that the corresponding proposition in the cloud is true at the world as A believes it to be.

One might worry that **Belief/Assertion** is in tension with the possibility of felicitously canceling the implication that the speaker has a certain conative attitude. (Recall the discussion of “cold” and “reverse” uses in Chapter 2). I think the tension is only apparent. The solution, I want to suggest, consists in specifying or complicating the *relation* between an agent and a standard in such a way that it allows for cold and reverse uses. A natural way to do so is to relax the acceptance-subscribing relation between a subject and a standard, for instance by allowing it to be *conditional*. More specifically, the idea is that we think of amoralists and immoralists as having a sort of conditional commitment to morality. When Jim says “eating fish is wrong, but I don't disapprove of

that she has the relevant attitude—seems different from the one mentioned in the text. (*Nota bene:* Copp opts for a conventional implicature account. But I don't think that's relevant to the present point).

³⁸What's more, one proposed explanation as to how the implicature is generated appeals precisely to an expectation that the speaker is sincere, where to be sincere amounts to having the relevant attitude.

³⁹Bedke (ms) outlines an account that goes in that direction. I briefly discuss his view below, in §3.5.

it” or “eating fish is wrong, but I approve of it”, he could be paraphrased as follows: “if I played the moral game, I would disapprove of eating fish. But I don’t care about this game” or: “If I played the moral game, I would disapprove of eating fish. But I prefer to play the *schmoral* game, where you do the exact opposite”. This would explain why cold and reverse utterances need *not* be incoherent. The appeal to a conditional commitment is *one way* of refining the acceptance relation. There may well be others. The general schema could be something like:

- **Belief/Assertion***: An agent *A* is entitled to believe/assert the cloud $|\varphi|^c$ only if *A* bears conative relation **R** to a standard s_i such that the corresponding proposition in the cloud is true at the world as *A* believes it to be.

Of course, this requires that *A* bears *some* conative relation to the standard in question. But I think this much is not especially controversial when *internal* uses of normative expressions are involved. Recall that cold/reverse uses are not the same as external uses.

There may be several ways of incorporating the envisaged explanation of cold and reverse uses in the pragmatic story given above. One is to say that moral utterances carry a generalized conversational implicature that the speaker possesses the unconditional attitude—this would be the default assumption given the practical import of moral discourse—, but this implication can be defeated precisely when the agent only possesses a conditional attitude, and cancels the implication to signal that. Alternatively, one could say that the implicature is that the speaker possesses *some* conative attitude or other, and constructions like “... but I don’t disapprove of *x*” could be taken to qualify/specify the nature of that attitude.

At this point, it will prove useful to return to the question of belief-ascriptions briefly touched on in §3.2.2. Earlier, I said that a sentence like “Mia believes that eating fish is wrong” communicates that Mia endorses/accepts a standard that forbids eating fish. Obviously, if my previous considerations are correct, this too will have to be modified in the way I suggested. Still, one might want to resist this sort of analysis for belief-ascriptions. For one might agree that something like **Belief/Assertion*** adequately

describes our practice, and hence the *correctness* conditions of judgments/assertions, but nonetheless think that no conative state whatsoever is *constitutive* of normative beliefs. The idea would be that for Jim to believe that eating fish is wrong is for Jim to believe that his standard forbids eating fish, where this doesn't require Jim to have a conative relation to that standard; the relation may be purely cognitive. If I understand him correctly, MacFarlane (2014, pp. 172-174) seems to suggest something along these lines in distinguishing his relativism from Gibbard's expressivism. The idea, adapted to the moral case, seems to be that to believe that eating fish is wrong is to believe that eating fish is wrong is true as assessed from one's standard/perspective. And while it would be irrational for one to be in that mental state without disapproving of eating fish, that's nonetheless a possible state one could be in.⁴⁰ Whether there is a significant difference between this proposal and the one I've sketched might turn on how the details are fleshed out. In any case, I'm leaving a more accurate treatment of these issues to future work.

3.4.2 Agreement and Disagreement

Let us now turn to the explanation of disagreement and agreement. Suppose that B endorses a standard that is incompatible with A's standard. A says: "eating fish is wrong". Why should B reply by saying "no, it's not wrong", "that's false", "I disagree"? Similarly, if A endorses a standard that is compatible with B's standard, how should we understand his use of expressions such as: "yes", "that's true", "I agree", and so on and so forth? What do these expressions target?

As far as I can see, there are at least three salient options:

- (i) The set of propositions, i.e. the semantic/asserted content.
- (ii) A subset of the propositions in the set.
- (iii) A pragmatically communicated content.⁴¹

⁴⁰Discussing the case of taste, MacFarlane (p. 174) writes: "Alex might find herself unable to believe that licorice is tasty because she aspires to greatness and thinks (on the basis of reading) that only uneducated people think that licorice is tasty. Though she likes the taste of licorice, and hence has sufficient reason to conclude that it tasty, she just can't bring herself to draw that conclusion. This is a form of irrationality, but perfectly intelligible."

⁴¹In principle, there are other possibilities: e.g. the sentence, the utterance, or the speech-act. However, at least when alethic assessments are involved (e.g. "that's true/false", "what you said is true/false")

Before I offer my explanation, it may be important to briefly refer to a distinction that is sometimes drawn in the literature between two sorts of normative disagreements: *non-fundamental* disagreements and *fundamental* disagreements. (See, e.g., Finlay 2014). It seems to me that the term “non-fundamental normative disagreement” is used in two slightly different senses that may be worth keeping separate. Sometimes it is used to refer to a disagreement that is ultimately grounded in a difference in non-normative beliefs (factual beliefs, or logical beliefs concerning what follows or doesn’t follow from a given set of rules). One might quibble over whether this is properly called a “normative” disagreement, but no matter. At other times, the term is used to refer to a disagreement that happens on the background of a significant agreement. So, it is sometimes said that the parties agree in their basic standards, but disagree on how their common standard applies to a certain case. In this second sense, a non-fundamental normative disagreement is plainly normative. Now, I’ve been working with a conception according to which a standard is *complete*, i.e. determines a classification of each action as required, permitted or forbidden. So, to capture the distinction at stake, I will say that the (incomplete) standard accepted by a group G is the standard that results by intersecting the several individual (complete) standards.

Notice that on this second gloss of the fundamental/non-fundamental distinction, the two notions appear to be gradable with the limiting case being the one where the overlapping area between the competing standards is almost empty. On the other hand, the first construal doesn’t allow for degrees (i.e. there isn’t a more or less fundamental disagreement). Obviously, the most problematic cases are those where the disagreement over whether, say, A *ought to* φ , is not grounded in a non-normative disagreement. These sorts of disagreements will be my main focus throughout this section.

Here is how the trivial descriptivist can account for disagreement. Suppose that A assertorically utters “eating fish is wrong” (1). She asserts the corresponding set of propositions \mathcal{P}_1 . Now, suppose that B ’s standard permits eating fish, and thus that it

and the like) as opposed to non-alethic ones (‘no’, ‘no, it’s not’ and the like) it seems odd to take a sentence/utterance or a speech-act as the entities of which truth and falsity are predicated. Talk of sentences/utterances/speech-acts as being true or false seems derivative on talking of their contents as being true or false.)

is incorrect for B to believe \mathcal{P}_1 , given that the propositions in the set that are about B's standard are false in all of B's belief worlds. It follows that B should object to A's claim, by asserting $\mathcal{P}_{(-1)}$, i.e. the set of propositions associated with "eating fish is not wrong" ($\neg 1$). Assuming that $\langle \text{it is false that } \mathcal{P}_{(i)} \rangle$ is equivalent to $\langle \neg \mathcal{P}_{(i)} \rangle$, which is in turn equivalent to $\langle \mathcal{P}_{(-i)} \rangle$, we can predict that B will also be entitled to respond to A's assertion by using constructions like: "that's false". In MacFarlane's vocabulary, we could say that A's and B's beliefs are *non-cotenable*: A/B cannot rationally accept both \mathcal{P}_1 and $\mathcal{P}_{(-1)}$. On this approach, the target of the that-clause is the assertoric content. There is a set of propositions such that A asserts that set and B asserts its negation (where to assert the negation of a set of propositions \mathcal{P} is to assert a set that contains the negation of each of the p_i in \mathcal{P}). Insofar as one is entitled to assert the negation of that content/set one is also entitled to say that it is false.⁴²

In light of these considerations, we can accommodate disagreement as follows:

- **Disagreement:** An agent B is entitled to reject A's assertion of $|\varphi|^c$ if he accepts a standard s_i such that the corresponding proposition in the cloud is false at the world as B believes it to be.

How does this explanation fit into the scoreboard model of the conversation? Simplifying, if B accepts A's assertion of "eating fish is wrong", i.e. agrees with it, then the information is added to the score that A and B are committed to standards that forbid eating fish. If B disagrees, the only information that is added is the one about the individual preferences of the two subjects. In order to better analyze the discourse dynamics of conversations about normative matters, we can enrich the structure of the conversational score and add a parameter that represents the conversationalists' mutually agreed-upon practical commitments—call it \mathcal{S} . Normative conversations indirectly aim at changing the value of this parameter.

One might think that the simple fact that A/B cannot accept what B/A believes-says without changing her mind or falling in state of inconsistency cannot fully explain why

⁴²A different option could be to take route (ii) and say that what is targeted by "that's false" is a *subset* of the propositions in \mathcal{P}_1 , namely those that correspond to the standards accepted by the disagreeing party B, which are indeed false in all of B's belief worlds.

A/B would engage in a dispute with B/A. For one might point out that subjects who believe and disbelieve the same world-neutral, time-neutral, or location-neutral proposition p (assuming, of course, that there are such propositions) while occupying distinct worlds, times or locations have non-cotenable beliefs, but have no reasons to criticize each other views. There is a plausible explanation, that many variabilists have toyed with, and that Carballo and Santorio (forthcoming) have recently given a fuller articulation.⁴³ Roughly, Carballo and Santorio suggest that when engaging in serious normative conversations, speakers presuppose *that they ought to converge on some norm or other*. This assumption is the counterpart of the presupposition, which regulates factual conversations, that there is an actual world on which speakers' beliefs ought to converge.⁴⁴ Crucially, these are *normative* presuppositions, and the *ought* in question is taken to capture “the normative force of Grice’s Cooperative Principle” (p. 2). The details need not concern us here. The important thing to notice is that, plausibly, there is no analogous presupposition in the case of, say, subjects that occupy distinct locations. If A, who lives in London, asserts: <it’s sunny> and B who lives in Rome asserts <it’s not sunny>, the rationality of their conversation (presumably on the phone) doesn’t require that there be a presupposition in place that they ought to converge on a single location.

3.4.3 Retraction

Let us finally turn to the case of *retraction*. A retraction, recall, is a speech-act which consists in taking back a prior speech-act in the light of a change in one’s normative views. Take the following example:

Jim-at-t₁: Capital punishment is not permissible.

Mia: You used to think and say otherwise when you were younger.

Jim-at-t₂: I was wrong, I take that back.

⁴³They do so in the context of explaining how communication works for metanormative expressivists, but I take it that the model can be adopted by other variabilists too.

⁴⁴The general idea is suggested, among others, by Gibbard (1990) and MacFarlane (2007), though at least the latter opts for a purely evolutionary-sociological explanation. See also Schafer (2014) for an explanation that is more in line with the one offered by Carballo and Santorio.

MacFarlane (2014) thinks that if Mia is right, then Jim-at- t_2 has an obligation to retract his previous assertion. This implies that it wouldn't be okay for him to say something like: "I know, I had different values at that time". It's not obvious to me that Jim-at- t_2 would be violating some linguistic/conceptual norm in refusing to retract along the lines suggested. I'm more inclined to think that he's *permitted*, but not required to retract. I think the data are delicate here, and absent some more decisive evidence I will stay neutral with respect to the permissibility/obligatoriness of retraction. At any rate, I think that a trivial descriptivist is free to choose between a pragmatics that requires, as opposed to one that merely permits, to retract one's previous normative assertions in the light of a change in one's normative perspective. On these grounds, I propose the following:

- **Retraction:** A-at- t_1 is entitled/required to retract A-at- t_2 's assertion of $|\varphi|^c$ if the members of $|\varphi|^c$ that are about the standard(s) accepted by A-at- t_2 are false at the world as A-at- t_2 believes it to be.

How can we understand this in the conversational score model? As MacFarlane (2011) points out, in a framework in which an assertion is a proposal, a retraction is best understood "not as the withdrawal of a proposal to change the context, but as a new proposal to undo the changes that were made" (p. 16). So, we can understand A-at- t_1 's retraction as an implicit proposal to make his previous standards invalid, i.e. to subtract them from the range of values of the score.

This concludes my explanation of the data at issue. In the next session, I compare my account to similar approaches advanced in the literature, and in the following I address some metasemantic questions.

3.5 Comparison With Similar Approaches

Trivial descriptivism bears some relevant resemblances with at least three views developed in the recent philosophical literature: von Fintel and Gillies' (2008, 2009) *cloudy contextualism* about epistemic modal claims, and Bedke's (ms) *GAP non-cognitivism*. I

will look at each of them in turn. Hopefully, this will also help get a better grasp of the proposed view.

A cloudy-contextualist about a sentence φ thinks that an utterance of φ is implicitly relative to a parameter (a body of information, a system of norms, etc.) determined by the context of utterance (this is what makes the view *contextualist*); unlike the ordinary contextualist, though, a cloudy-contextualist thinks that there may be a significant degree of *contextual indeterminacy*, i.e. a number of ways of resolving the parameter that the sentence is semantically associated with—or, in von Fintel and Gillies’ metaphor, “a cloud of admissible contexts” with respect to which the sentence can be interpreted (this is what makes the view *cloudy*). For instance, if we take an utterance of “might φ ”, a speaker A, and a hearer B, the the admissible interpretations will be, at least: A’s body of information is compatible with φ ; B’s body of information is compatible with φ ; A+B’s body of information is compatible with φ .

Von Fintel and Gillies (hereafter, vF&G) couple their semantics with a special pragmatics. Utterances of epistemic modal sentences like “The treasure might be under the palm” are *sui generis* speech-acts which consist in “putting into play” or “floating” a cloud of propositions, one for each possible resolution of the “body of information”-parameter that the sentence is semantically associated with. As a first approximation, to “put into play” many propositions is to do something akin to introducing them into the conversation, as available for consequent assessments of truth and falsity.

vF&G offer the following story to explain the various linguistic data concerning assertion, disagreement, and retraction. A speaker A is justified in making the utterance just in case he’s justified in asserting at least one of the propositions put-into-play, plausibly the solipsistic proposition (\langle The treasure might be under the palm given A’s body of information \rangle); a hearer B can deny the utterance just in case he’s justified in denying the strongest proposition put-into-play, e.g. the group-proposition (\langle The treasure might be under the palm given A+B’s body of information \rangle). Likewise, a speaker A can retract his previous utterance just in case he’s justified in denying the strongest proposition put-into-play (\langle The treasure might be under the palm given A’s body of information at

t*>).

How does trivial descriptivism differ from cloudy contextualism? There is a difference in the domain of application of the view, a couple of differences in the semantics, and a *possible* difference in the pragmatics. The first obvious contrast is that cloudy contextualism is offered to handle epistemic modals, whereas trivial descriptivism is a theory of normative language, modalized and otherwise. I think that vF&G’s approach can be extended to normative discourse, but of course this involves some adjustment (more on this below). The second contrast lies in the fact that cloudy contextualism is semantically more conservative than trivial descriptivism. vF&G leaves the standard contextualist formal semantics untouched (each sentence-in-context is assigned a unique proposition as its semantic value), while the trivial descriptivist introduces a new view about the very semantic contents of sentences. (I will say a bit more about the semantics adopted by vF&G in §3.7). The third difference consists in the fact that trivial descriptivism is not a contextualist view, in that it doesn’t hold that the content of normative claims depends on features of the context of use. Although vF&G include among the “admissible contexts” the state of information of people who are not among the participants to the conversation (so-called “eavesdroppers”), they seem to posit some constraints (presumably fixed by factors of the context of use) on the set of admissible contexts. (After all, if they didn’t put any constraints on the range of admissible contexts, “contextualism” would be a bit of a misnomer for their proposal.)⁴⁵ The fourth and last point concerns vF&G’s talk of “putting in play” or “floating” a proposition. Above, I assumed that these acts are to be construed as *illocutionary* acts. This is perhaps not the unique way to understand their view. So, I will now sketch two alternative interpretations, and relate them to the trivial descriptivist proposal.

As far as I can see, there are two ways of understanding the talk of “putting in play”:

- (i) To put into play a proposition p is to utter a sentence that *semantically expresses* p . Consequently, to put into play a set of propositions $\{p_1, p_2, \dots, p_n\}$ is to utter a

⁴⁵Note that this is different from saying that if one thinks that the context (i.e., in the case under consideration, the domain of quantification of the modal) is, say, the state of information of everyone in the world, one is not a contextualist anymore. If to be a contextualist is to think that content depends on context, then this theorist has a contextualist semantics, a universalist contextualist semantics if you wish)

sentence that semantically expresses: $\{p_1, p_2, \dots, p_n\}$. We can perhaps characterize this speech-act by saying that it is a *locutionary* act in Austin's sense, where to perform this act is "roughly equivalent to uttering a certain sentence with a certain sense and reference" (Austin 1962, p. 108).

- (ii) To put into play a proposition p is to make an illocutionary speech-act of some kind, and not just to express the semantic content of p .

I take it that option (i) is incomplete as it stands. Something needs to be said about the illocutionary force of the speech-act. Arguably, the most natural option is to say that semantic content and asserted content can diverge in theoretically interesting ways; in the case of our concern, only some of the propositions semantically expressed would be asserted.⁴⁶ Two considerations seem to count against this reading of cloudy contextualism. The first is that it is unclear how the mere fact of being semantically expressed (where this is something devoid of any illocutionary force) can be sufficient to make a proposition available to be evaluated for truth and falsity. Compare, perhaps, ordinary cases of ambiguity. We might think that all the disambiguations are semantically expressed; but something more seems required to get the result that all the disambiguations become candidates for acceptance/rejection by a cooperative hearer. The second consideration is the following passage, which seems to strongly support hypothesis (ii):

Alex is justified in uttering the BEM [bare epistemic modal] because she is in a position to flat out assert the *A*-reading. But just asserting the *A*-reading isn't what she's doing. When Alex puts the three propositions in play, the other readings, the hearer-centric reading and the strong group reading, are floated. She does not have to be in a position to assert those in order for them to be available for Billy to react to. It is as if she is conjecturing that the *B*-reading and the *A+B*-reading are true or asking whether they are true. (vF&G 2009, p. 121).

The passage seems to suggest a "multiple speech acts" analysis. By uttering a bare epistemic modal, Alex performs two speech-acts: she *asserts* one proposition (the *A*-reading) and she *conjectures* that the other two propositions (the *B*-reading and *A+B* reading) are true (or she *asks* whether they are true).⁴⁷

⁴⁶The distinction between semantic content and asserted content plays an important role in Bach (1994, 2001) and Soames 2005, 2009).

⁴⁷This reading seems confirmed by the fact that vF&G entertain the possibility of such account in their (2007).

If this is the right interpretation of cloudy contextualism, then I think the fourth difference with trivial descriptivism partly evaporates. Both views countenance an indirect speech-act in addition to the assertoric speech act. To be sure, the following two differences remain. First, the cloudy contextualist says that some propositions are literally asserted; I take this to mean that they are added to the conversational score in the usual way, i.e. directly. As we have seen, this is not the case for the trivial descriptivist: the literal content of a normative utterance has no non-trivial effect on the score. Second, as one would expect, the nature of the indirect speech act is somewhat different in the two cases: if in the epistemic case, one *asks/conjectures* whether some p_i is true, or highlight a certain possibility, in the deontic case one *proposes* that some p_i be true, or expresses his endorsement of the standard that would verify p_i .

Perhaps, the relevant difference boils down to one's favorite way of classifying speech-acts. One might raise the following objection to the trivial descriptivist. The view says that (unembedded) normative sentences are used to make assertions, but that the purpose of these assertions is to make the world fit the mind, rather than to make the mind fit the world. However, if this is the right "direction of fit", what's the point of calling these speech-acts assertions? Isn't it definitional of assertoric uses of language that their aim is to represent the world accurately, or at least that they are regulated by some truth-norm? Perhaps, so the worry goes, trivial descriptivism is better construed as the view that we should introduce a new category of speech-acts, or at least understand normative utterances as speech-acts other than assertions.

The answer to this question, I think, turns on how one individuates speech-acts in general, and assertoric speech-acts in particular. Focusing on the latter, we can distinguish between four views: (A1) assertions as expressions of attitudes, e.g. beliefs; (A2) assertions as proposals to change the conversational score by set intersection; (A3) assertions as moves constitutively governed by some specific norms, e.g. a truth-norm or a knowledge-norm; (A4) assertions as the undertaking of certain commitments or obligations.⁴⁸ Utterances of unembedded normative sentences count as assertions if we adopt

⁴⁸This taxonomy roughly follows MacFarlane (2011) and Cappelen (2011).

(A2) or, modulo some qualifications, (A1). However, if one thinks that assertions are constitutively governed by a truth or knowledge norm, then I think one will be forced to adopt a story along the lines of vF&G, or to drop the assumption that those speech-acts are/involve assertions. The upshot, as I see it, is that the choice depends on one's preferred taxonomy for speech-acts. Should there be compelling reasons for preferring a categorizations like (A3), I think the trivial descriptivist could simply drop the idea that utterances of normative sentences are assertions. I don't think this would make the view less attractive.

If we consider the normative case, a very similar account to the one proposed here is developed by Bedke (ms), under the heading of "GAP non-cognitivism". Bedke summarizes GAP non-cognitivism as the conjunction of the following two claims: "i) thin normative terms are maximally vague with respect to descriptive semantics, and ii) they have non-descriptive use conditions fixed by *when* use rules that are sensitive to the non-cognitive attitudes of the speaker". The first feature roughly corresponds to the semantic hypothesis I have myself articulated. (I will highlight some differences in a moment). The second feature builds on a distinction between two kinds of use rules, that Bedke calls "*for* rules" and "*when* rules" respectively. Use rules are just what you would think: rules concerning the correctness/incorrectness of uses of bits of language. They correspond to what I've been calling the *pragmatics* in the general scheme of a theory of meaning. The basic idea is that we use normative terms *when* certain conditions obtain, namely conditions picked out by our non-cognitive attitudes, but we don't use them *for* these conditions, as it happens instead with descriptive terms that are used for the relevant conditions, and hence describe their content.

How does GAP non-cognitivism differ from trivial descriptivism? I think there are two differences, but as I am going to explain they appear to be relatively minor. The first is that, in Bedke's proposal, the maximal vagueness hypothesis is implemented by drawing from existing formal frameworks devised to assign extensions to vague sentences. In particular, Bedke remains neutral with respect to the possibility of adopting a supervaluationist semantics, or a three-valued one, while I've instead used a relativist semantics

to assign extensions to complex sentences (a semantics that basically corresponds to supervaluationism without the identification of truth with supertruth). More importantly, as I understand him, Bedke doesn't commit himself to any view about the *propositional* contents of vague and normative sentences, i.e. the question of which proposition, if any, is expressed by vague and normative sentences (e.g. one, many, none, it's indeterminate). Relatedly, he remains open to the possibility of treating the case of normative thought differently from that of normative language (pp. 32-33, fn. 30).

A second possible difference concerns point (ii). Bedke considers but eventually discards the possibility of a view on which normative terms are maximally vague (this is point (i)), but attitudes are only pragmatically conveyed, where "pragmatically" here means that they are not part of what is conventionally communicated. His rationale for this choice is twofold. First, he doubts that canceling the inference that one has the conative attitude can ever be felicitous, when the term is used in a normative sense. I've committed myself otherwise above. However, I'm not sure that the disagreement is, in fact, substantive. For instance, Bedke consider someone who says: "Tormenting the cat is wrong, but I have nothing against it" and observes that: "S/he is using the wrong term to render a judgment about tormenting the cat if s/he lacks *any* relevant negative attitude" (p. 22). I agree, in a sense. I think the key is in the "any" (his italics), and in in the way in which one understands what it is to have a pro-attitude, or to subscribe to a standard. Second, he thinks that a merely pragmatic story leaves the case of normative *thought* unexplained. Again, I agree, and I've tried to address this worry above. To be sure, as mentioned above, Bedke doesn't appeal to implicatures, but rather gives an account according to which, roughly, the use of predicate like "is wrong" is governed by a *when* use rule, of the following kind: "Predicate 'is wrong' of a situated action, A, when A exhibits dthat(a property the speaker all in disapproves of in actions in such situations)" (p. 25). So, while the descriptive *for* use rules leave a "maximal *gap of optionality* when it comes to use" (p. 27), *when* use rules fill the gap. In advancing this proposal, Bedke draws from what Gutzmann (2015) calls a "use-conditional semantics", a formal approach inspired in turn by Kaplan's (1999, not formally published) "The Meaning of Ouch and Oops". I

don't have much to say about this, except that it seems a promising avenue for the kind of non-cognitivist alternatives to expressivism that both Bedke and I are interested in exploring.

3.6 Metasemantic Issues

Plausibly, it is not a brute fact that normative expressions have the meanings that they do. In the case at hand, what metasemantic view can deliver the semantic facts posited by trivial descriptivism? That is, what non-semantic facts (i.e. facts about speakers' intentions, conventions, causal connections, etc.) *ground* or *explain* the particular semantic facts described by the proposed account?

Clearly, answering this question in a satisfactory way is a tall order. My aim here is to simply sketch some possible replies. In particular, I will offer three models one could adopt.

Here is one first candidate. Start with a neo-descriptivist metasemantics of the kind outlined by Lewis (1997) and Jackson (1998). Roughly, according to this picture there is a set of widely shared (implicit) beliefs about normative issues—a so-called “folk theory” about normative matters—and the extension of a normative term ε is determined by what best satisfies such theory. Now, according to the trivial descriptivist moral terms lack an extension, insofar as they lack a unique content: their character is a function from a context to a set of contents/concepts. What kind of descriptions do speakers associate with moral terms so that they end up being like this? There may be various options. We could say that the platitudes surrounding moral vocabulary are too *thin* and *vague* to allow us to pin down a unique extension. Consider, for instance, the apparent platitude that moral considerations must bear some relation to the the good and harm of other sentient beings.⁴⁹ Even if we think that this is a platitude, it doesn't seem to rule out any moral system. (The sociopath's well-being might well depend on his performing acts of gratuitous cruelty). A different suggestion could go as follows. We could say that it is written into our folk morality that there is no particular descriptive property

⁴⁹This is a more liberal version of Foot's conceptual constraint, which is restricted to humans.

that deserves to be called “the” property of being right, or wrong, or good, etc. This is, in effect, what Yablo suggests in his review of Jackson’s (1998). He criticizes Jackson’s contention that “it is part of current folk morality that convergence will or would occur” (p. 18) by submitting that “a certain kind of anti-dogmatism” is instead a condition of moral communication. That is, it is part of our conception of matters normative that “nothing will ever be regarded as the point of arrival: the point at which reference is finally fixed and moral theory acquires a conceptual imprimatur”. According to Yablo’s version of the folk theory of morality “our primary commitment [...] is that ‘right’ and similar action-guiding terms should stand for the same or similar properties in all our mouths”. For this reason “we refuse to attach ‘right’ to any particular property so tightly that moral dissidents, even hypothetical ones, come out as simply misusing ‘right’” (p. 19). Clearly, both stories are compatible with there being other platitudes, for instance that normative concepts have some defeasible connection with motivation, that no two actions can differ normatively without also differing in their natural features, and so on.

Perhaps a descriptivist metasemantics doesn’t sit well with the trivial descriptivist claim that normative sentences lack a descriptive content, i.e. don’t rule in and out any possible world. One might think that expressivist or non-cognitivist analyses may be a more natural place to look for a satisfactory account. Recently, many authors have suggested that expressivism is best cast not as a thesis about the *linguistic meaning* of the target sentences, but rather (i) as a thesis about the *use* of these sentences (this is the strategy that Yalcin (2012) seems to favor)⁵⁰ or alternatively (ii) as a thesis in *metasemantics*, rather than first-order semantics (as Ridge 2014 proposes). Consider the first route. We could say that the fact that normative terms, unlike purely expressive words such as “ouch” function as predicates, that the sentences that feature them can be embedded in truth-functional and non truth-functional contexts, and so on and so forth, suggest that we treat them in a way that mirrors terms that obey a descriptivist (i.e. cognitivist) semantics. However, the *use* of these sentences is, roughly, as the expressivist

⁵⁰Yalcin submits that the expressivist idea is “is best developed as a thesis in formal pragmatics—specifically, as a thesis about the characteristic sort of update to the conversational common ground a normative or plan-expressing sentence serves to trigger” (pp. 11-12).

says: namely, they are used to express and promote non-cognitive attitudes. If we add to this the fact that they can be used to express any attitude whatsoever, this would account for the fact that normative predicates lack an extension, and normative utterances lack truth-values. One could perhaps try to achieve the same result by adopting an expressivist metasemantics, according to which normative predicates have the contents that they do in virtue of expressing some non-cognitive attitude.

You might think that trivial descriptivism is surely incompatible with *externalist* metasemantics, for instance with a natural kind semantics for normative terms. Now, this would not be, by itself, a major problem. A battery of thought experiments, ranging from Hare's cannibals and missionaries to Horgan and Timmons' moral twin earth, provide good evidence that normative terms, unlike terms designating, say, biological or chemical kinds, are non-twin-earthable. But forget about that. Let's grant that normative terms function *pretty much like* natural kind terms. This doesn't yet rule out the possibility that these terms turn out to be radically indeterminate. Consider Kripke's take on unicorns (see Kripke 1980 and 2013). Very roughly, he thought (i) that "unicorn" is a natural kind term; (ii) that the myth didn't say enough to pick out a unique species; and therefore (iii) that "no counterfactual situation is properly describable as one in which there would have been unicorns" (Kripke 1980, p. 156). It is not hard to see how, *mutatis mutandis*, a similar case could be made for normative terms.

Interestingly, Horgan and Timmons (2000) think that some versions of a natural kind semantics for normative terms⁵¹ might have precisely this consequence, insofar as generic notions like that of "flourishing" or "impartiality" do not secure a reference for normative terms (ibid., p. 45). Of course, they take all this to provide a *reductio* of the view under consideration. And rightly so: surely, a realist didn't want *that* from a natural kind semantics. But all this shows how the trivial descriptivist hypothesis could in principle be coupled also with an externalist picture of meaning.

There is a further question that is in some sense metasemantic, and that one might ask, namely: why did we introduce concepts that work in this way? This question

⁵¹Like the one developed by Copp (...) according to which "right" is more similar to a term like "milk" or "honey", than to a term like "water". More precisely, he puts forward a view according to which [...].

asks for some rationalizing explanation for the existence of some linguistic conventions, rather than for some reductive or grounding explanation. Here is a sketch of one such explanation. The crucial role of these terms, let's say, is to present some features in a positive or negative light. Plausibly, each of us has a moral and aesthetic ideal, i.e. a picture of what he would like the world morally or aesthetically to be. Communicating and discussing this picture is important in various respects, mostly in the context of planning and coordinating actions. Arguably, this goal is best achieved by a vocabulary that, first (i) is not tied to any specific set of descriptive features, and is thus flexible, negotiable, and so on; and second (ii) that allows us to talk about our desires and preferences, without referring to them (or their contents) or describing them (or their contents). I think that a language that works as the trivial descriptivist suggests is able to deliver precisely this result.

3.7 Keeping the Cloud Out Of The Semantics

I have chosen to frame my account in terms of a cloudy semantics—i.e. a semantics that assigns sets or clouds of propositions as semantic values of sentences-in-context. However, this is not the only way of modeling the hypothesis that an utterance of a sentence is associated with a set of propositions. One could opt for a cloudy pragmatics, according to which each sentence-context pair $\langle s, c \rangle$ expresses a unique proposition p , but for any given concrete utterance u there is a cloud of abstract sentence-context pairs $\langle s, c \rangle_1, \langle s, c \rangle_2, \dots, \langle s, c \rangle_3$ each of which correctly models or represents the concrete utterance u .⁵² Since there are two formal objects—namely, a sentence and a context—this indeterminacy can in turn come either from the sentence or from the context. In the first case, it is indeterminate which base-language (i.e. disambiguated) sentence corresponds to a given surface-language sentence—call this “presemantic indeterminacy”; alternatively, it is indeterminate which character is associated with a certain string of sounds or marks—

⁵²Note that here I'm using the adjective “pragmatic” in a very narrow sense to refer, roughly, to what is not part of the Kaplanian character. The term “pragmatics” is more often used for the theory of speech acts (for what I discussed in §3.3) and also for the Gricean theory of implicatures.

call this “character-indeterminacy”.⁵³ In the second, it is indeterminate which abstract context (i.e. which ordered n -tuple of values) corresponds to, or correctly maps, a given concrete speech-context—call this “context-indeterminacy”.

In a nutshell, the difference is that while a cloudy semantics introduces a new semantic relation (one sentence-in-context/many propositions), a cloudy pragmatics (of any of the three varieties just distinguished) sticks to the standard one-to-one relation (one sentence-in-context/one proposition), but says that it is indeterminate to which proposition in a certain set of propositions the classical/familiar semantic relation holds.

As far as I can tell, one can cast the basic idea behind trivial descriptivism in these alternative frameworks too. One noteworthy example of a cloudy pragmatics is the cloudy contextualism of von Fintel and Gillies, discussed above. In the terminology introduced above, their view posits a form of context-indeterminacy.

I think that the choice between a cloudy semantics and a more conservative cloudy pragmatics may ultimately turn on overall cost-benefit considerations, having to do with explanatory power, comparative simplicity, unification, and so on. I would also like to point out that the possibility of more conservative implementations of the cloudy hypothesis should make the view more palatable to those who are not willing to build the cloudiness into the semantics, but who at the same time prefer the orthodox static, truth-conditional, and non-relativist semantics to the new frameworks advanced in the literature.

⁵³The last option corresponds to Lewis’s idea, in “Languages and Language”, that linguistic expressions may be associated with a cluster of characters, because speakers’ behaviour is compatible with them using several different formally precise languages. In other words, there is range of possible (classical, precise, etc.) languages, and it is underdetermined which language is the one actually spoken by a certain population (cf. Lewis 1975, pp. xx).

Bibliography

- Ayer, A. J. (1952), *Language, Truth and Logic*: Dover Publications.
- Baker, D. and J. Woods (2015), 'How Expressivists Can and Should Explain Inconsistency', *Ethics* **125**, No. 2, 391–424.
- Balaguer, M. (2011), 'Bare Bones Moral Realism and the Objections from Relativism', Edited by Hales, S., *A Companion To Relativism*: Wiley-Blackwell.
- Beall, J. C. (2010), 'Vague Intensions: A Modest Marriage Proposal', Edited by Dietz, R. and S. Moruzzi, *Cuts and Clouds*: Oxford Scholarship Online.
- Björklund, F., G. Björnsson, J. Eriksson, R. Francén, and C. Strandberg (2012), 'Recent Work on Motivational Internalism', *Analysis* **72**, No. 1, 124–137.
- Björnsson, G. and S. Finlay (2010), 'Metaethical Contextualism Defended', *Ethics* **121**, No. 1, 7–36.
- Blackburn, S. (1993), *Essays in Quasi-Realism*: Oxford University Press.
- Boghossian, P. (1990), 'The Status of Content', *Philosophical Review* **99**, No. 2, 157–84.
- (2006), 'What is Relativism?', Edited by Patrick Greenough and Michael Lynch, *Truth and Relativism*: Clarendon Press.
- Braun, D. and T. Sider (2007), 'Vague, So Untrue', *Noûs* **41**, No. 2, 133–156.
- Brink, D. O. (1989), *Moral Realism and the Foundations of Ethics*: Cambridge University Press.
- Brogaard, B. (2008), 'Moral Contextualism and Moral Relativism', *Philosophical Quarterly* **58**, No. 232, 385–409.
- Cappelen, H. and E. Lepore (2005), *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*: Blackwell Pub., 2005.
- Cappelen, H. (2011), 'Against Assertion', Edited by Brown, J. and H. Cappelen, *Assertion: New Philosophical Essays*: Oxford University Press.
- Carnap, R. (1947), *Meaning and Necessity*: University of Chicago Press.

- Carston, R. (1996), 'Metalinguistic Negation and Echoic Use', *Journal of Pragmatics* **25**.
- Chierchia, G. and McConnell-Ginet, S. (2000), *Meaning and Grammar: An Introduction to Semantics*: MIT Press.
- Copp, D. (1995), *Morality, Normativity, and Society*: Oxford University Press.
- (2001), 'Realist-Expressivism: A Neglected Option for Moral Realism', *Social Philosophy and Policy* **18**, No. 02, 1–43.
- Cuneo, T. and R. Shafer-Landau (2014), 'The Moral Fixed Points: New Directions for Moral Nonnaturalism', *Philosophical Studies* **171**, No. 3, 399–443.
- Dancy, J. (2004), *Ethics Without Principles*: Oxford University Press.
- Davis, J. (2015), 'Faultless Disagreement, Cognitive Command, and Epistemic Peers', *Synthese* **192**, No. 1, 1–24.
- Dowell, J. (2013), 'Flexible Contextualism About Deontic Modals: A Puzzle About Information-Sensitivity', *Inquiry* **56**, No. 2-3, 149–178.
- Dreier, J. (1990), 'Internalism and Speaker Relativism', *Ethics* **101**, No. 1, 6–26.
- Egan, A. (2010), 'Disputing About Taste', Edited by Warfield, T. and R. Feldman, *Disagreement*: Oxford University Press.
- Einheuser, I. (2008), 'Three forms of truth relativism', Edited by Carpintero, G. and M. Kölbel, *Relative Truth*, Oxford: Oxford University Press.
- Enoch, D. (2011), *Taking Morality Seriously: A Defense of Robust Realism*: Oxford University Press.
- Evers, D. (2011), 'The Standard-Relational Theory of 'Ought' and the Oughtistic Theory of Reasons', *Australasian Journal of Philosophy* **89**, No. 1, 131–147.
- Field, H. (2009), 'Epistemology Without Metaphysics', *Philosophical Studies* **143**, No. 2, 249–290.
- Fine, K. (1975), 'Vagueness, Truth and Logic', *Synthese* **30**, No. 3-4, 265–300.
- Finlay, S. (2005), 'Value and Implicature', *Philosophers' Imprint* **5**, No. 4, 1–20.
- (2009), 'Oughts and Ends', *Philosophical Studies* **143**, No. 3, 315–340.
- (2014), *Confusion of Tongues: A Theory of Normative Language*: Oxford University Press.
- von Fintel, K. and A. Gillies (2008), 'CIA Leaks', *Philosophical Review* **117**, No. 1, 77–98.
- (2009), "'Might' Made Right", Edited by Egan, E. and B. Weatherson, *Epistemic Modality*, Oxford: Oxford University Press.
- Foot, P. (1978), *Virtues and Vices and Other Essays in Moral Philosophy*: Oxford University Press.

- Geach, P. (1982), 'Whatever Happened to Deontic Logic?', *Philosophia* **11**, No. 1-2, 1–12.
- Gibbard, A. (1990), *Wise Choices, Apt Feelings*, Cambridge MA: Oxford University Press.
- (2003), *Thinking How to Live*, Cambridge MA: Harvard University Press.
- (2003), 'Normative Properties', *The Southern Journal of Philosophy* **41**, No. S1, 141–157.
- Gillon, B. (1990), 'Ambiguity, generality, and indeterminacy: Tests and definitions', *Synthese* **85**, 391–416.
- Glanzberg, M. (2015), 'Logical Consequence and Natural Language', Edited by Caret, C. and O. Hjortland, *Foundations of Logical Consequence*: Oxford University Press.
- Goodman, N. (1949), 'On Likeness of Meaning', *Analysis*, 1–7.
- Hare, R. M. (1952), *The Language of Morals*: Oxford Clarendon Press.
- Harman, G. (1975), 'Moral Relativism Defended', *Philosophical Review* **84**, No. 1, 3–22.
- (2015), 'Moral Relativism is Moral Realism', *Philosophical Studies* **172**, No. 4, 855–863.
- Horn, L. (1989), *A Natural History of Negation*: University of Chicago Press.
- Jackson, F., G. Oppy, and M. Smith (1994), 'Minimalism and Truth Aptness', *Mind* **103**, No. 411, 287–302.
- Jackson, F. (1998), *From Metaphysics to Ethics: A Defence of Conceptual Analysis*: Oxford University Press.
- Jackson, F. and P. Pettit (1998), 'A Problem for Expressivism', *Analysis* **58**, No. 4, 239–251.
- Kalderon, M. E. (2005), *Moral Fictionalism*: Oxford University Press.
- Kaplan, D. (1977), 'Demonstratives', Edited by Joseph Almog and John Perry and Howard Wettstein, *Themes From Kaplan*: Oxford University Press.
- King, J. (2014), 'The Metasemantics of Contextual Sensitivity', Edited by Burgess, A. and B. Sherman, *Metasemantics: New Essays on the Foundations of Meaning*: Oxford University Press.
- Kölbel, M. (2004), 'Faultless Disagreement', *Proceedings of the Aristotelian Society* **104**, No. 1, 53–73.
- (2009), 'The Evidence for Relativism', *Synthese* **166**, No. 2, 375–395.
- (2015), 'Relativism 2: Semantic Content', *Philosophy Compass* **10**, No. 1, 52–67.
- Korsgaard, C. (1996), *The Sources of Normativity*: Cambridge University Press.
- Kratzer, A. (1977), 'What 'Must' and 'Can' Must and Can Mean', *Linguistics and Philosophy* **1**, No. 3, 337–355.

- Kratzer, Angelika (1981), 'The Notional Category of Modality', Edited by H. J. Eikmeyer and H. Rieser, *Words, worlds, and contexts: New approaches in word semantics*: de Gruyter, Berlin.
- Lewis, D. (1970), 'General Semantics', *Synthese* **22**, No. 1-2, 18–67.
- (1980), 'Index, Context, and Content', Edited by Stig Kanger and Sven Öhman, *Philosophy and Grammar*: Reidel.
- (1989)', *Proceedings of the Aristotelian Society* **63**, 113–137.
- MacFarlane, J. (2007), 'Relativism and Disagreement', *Philosophical Studies* **132**, No. 1, 17–31.
- (2011), 'What is Assertion?', Edited by Brown, J. and H. Cappelen, *Assertion: New Philosophical Essays*: Oxford University Press.
- (2014), *Assessment Sensitivity: Relative Truth and its Applications*, Oxford: Clarendon Press.
- Mackie, J. L. (1977), *Ethics: Inventing Right and Wrong*: Penguin.
- McKenna, R. (2014), 'Disagreeing About 'Ought'', *Ethics* **124**, No. 3, 589–597.
- Moore, G. E. (1903), *Principia Ethica*: Dover Publications.
- Phillips, D. (1997), 'How to Be a Moral Relativist', *Southern Journal of Philosophy* **35**, No. 3, 393–417.
- Plunkett, D. and T. Sundell (2013), 'Disagreement and the Semantics of Normative and Evaluative Terms', *Philosophers' Imprint* **13**, No. 23.
- Poesio, M. (1996), 'Semantic ambiguity and perceived ambiguity', Edited by van Deemter, K. and S. Peters, *Semantic Ambiguity and Underspecification*: Stanford, CA: CSLI.
- Ridge, M. (2014), *Impassioned Belief*: OUP Oxford, 2014.
- Rovane, C. (2012), 'How to Formulate Relativism', Edited by Wright, C. and A. Coliva, *Mind, Meaning, and Knowledge: Themes From the Philosophy of Crispin Wright*: Oxford University Press.
- Schaffer, J. (2012), 'Necessitarian Propositions', *Synthese* **189**, No. 1, 119–162.
- Schiffer, S. (2003), *The Things We Mean*, Oxford: Oxford University Press.
- Schlenker, P. (2003), 'A Plea for Monsters', *Linguistics and Philosophy* **26**, No. 1, 29–120.
- Schroeder, M. (2008), *Being For*: Oxford University Press.
- (2008), 'Expression for Expressivists', *Philosophy and Phenomenological Research* **76**, No. 1, 86–116.
- (2014), *Explaining the Reasons We Share: Explanation and Expression in Ethics, Volume 1*: Oxford University Press.

- (2015), *Expressing Our Attitudes: Explanation and Expression in Ethics, Volume 2*: Oxford University Press.
- Shafer-Landau, R. (1994), 'Ethical Disagreement, Ethical Objectivism and Moral Indeterminacy', *Philosophy and Phenomenological Research* **54**, No. 2, 331–344.
- (1995), 'Vagueness, Borderline Cases and Moral Realism', *American Philosophical Quarterly* **32**, No. 1, 83–96.
- (2003), *Moral Realism: A Defence*: Oxford University Press.
- Silk, A. (2015), 'How to Be an Ethical Expressivist', *Philosophy and Phenomenological Research* **91**, No. 1, 47–81.
- (forthcoming), 'Normative Language in Context', *Oxford Studies in Metaethics* **12**.
- Smith, M. (1994), *The Moral Problem*: Blackwell.
- Soames, S. (1989), 'Semantics and Semantic Competence', *Philosophical Perspectives* **3**, 575–596.
- (2008), 'Drawing the Line Between Meaning and Implicature—and Relating both to Assertion', *Nous* **42**, No. 3, 529–554.
- Spencer, J. (2014), 'Relativity and Degrees of Relationality', *Philosophy and Phenomenological Research*, 1–28.
- Strandberg, C. (2012), 'A Dual Aspect Account of Moral Language', *Philosophy and Phenomenological Research* **84**, No. 1, 87–122.
- Street, S. (2009), 'In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters', *Philosophical Issues* **19**, No. 1, 273–298.
- Streumer, B. (2008), 'Are There Irreducibly Normative Properties?', *Australasian Journal of Philosophy* **86**, No. 4, 537–561.
- Sundell, T. (2011), 'Disagreements About Taste', *Philosophical Studies* **155**, No. 2, 267–288.
- Toppinen, T. (2015), 'Pure Expressivism and Motivational Internalism', Edited by Björklund, F., G. Björnsson, J. Eriksson, R. Francén, and C. Strandberg, *Motivational Internalism*: Oxford University Press.
- Unger, P. (1979), 'There Are No Ordinary Things', *Synthese* **41**, No. 2, 117–154.
- Wedgwood, R. (2001), 'Conceptual Role Semantics for Moral Terms', *Philosophical Review* **110**, No. 1, 1–30.
- (2010), 'The Moral Evil Demons', Edited by R. Feldman and T. Warfield, *Disagreement*: Oxford University Press.

- Wong, D. (1984), *Moral Relativity*. University of California Press, Berkeley.
- Woods, J. (2014), 'Expressivism and Moore's Paradox', *Philosophers' Imprint* **14**, No. 5, 1–12.
- Wright, C. (2001), 'On Being in a Quandary', *Mind* **110**, 45–98.
- (2006), 'Intuitionism, Realism, Relativism and Rhubarb', Edited by Greenough, P. and M. Lynch, *Truth and Realism*: Clarendon Press.
- (2008), 'Relativism about Truth Itself: Haphazard Thoughts about the Very Idea', Edited by Carpintero, G. and M. K'olbel, *Relative Truth*, Oxford: Oxford University Press.
- Yalcin, S. (2011), 'Nonfactualism About Epistemic Modality', Edited by Andy Egan and B. Weatherson, *Epistemic Modality*: Oxford University Press.
- (2012), 'Comments on Mark Schroeder, Being For', *Central APA*.
- (2014), 'Semantics and Metasemantics in the Context of Generative Grammar', Edited by A. Burgess and B. Sherman, *Metasemantics: New Essays on the Foundations of Meaning*: Oxford University Press.