8

Thick Concepts and Underdetermination

Pekka Väyrynen

8.1 Introduction

What is supposed to be distinctive of "thick" concepts relative to "thin" concepts in ethics and aesthetics is that thick concepts somehow "hold together" evaluation and non-evaluative description, whereas thin concepts are somehow more purely evaluative or normative.[1] This seems to capture an intuitive contrast between thick concepts, such as those expressed by terms like *rude*, *brutal*, *graceful*, and *kind*, and thin concepts, such as many concepts expressed by terms like *wrong*, *good*, and *impermissible*. Even if causing offense is both rude and bad, only *rude* seems to require as a matter of meaning that things falling under it must have something to do with causing offense; the meaning of *bad* generates no such constraint, since many bad things have nothing to do with causing offense. Thick terms and concepts seem to bear some broadly conceptual connections to some fairly concrete sort of non-evaluative descriptions.

What is more controversial is whether thick terms and concepts bear similar conceptual relations to evaluation. (By "evaluations" I mean claims or information with evaluative content, not mental acts of evaluation.) No doubt *rude* is typically used to convey negative evaluation, where "convey" is an umbrella term for different means (such as content, presupposition, or implicature) by which utterances can transfer information. Just what kind of failure of understanding would be manifested by someone who has caused offense but fails to grasp the kind of negative evaluation that is typically conveyed by calling something *rude* depends precisely on just how thick terms and concepts are related to the evaluations they may be used to convey. The standard view is that evaluation is built into the meaning (sense, semantic content) of utterances involving thick terms and concepts as much as the non-evaluative descriptions that they entail are so built. The alternative is that evaluations are some weaker, perhaps broadly, pragmatic

---

[1] For discussion of whether the distinction between thick and thin epistemic concepts is analogous, see Väyrynen (2008).

implication of utterances involving thick terms and concepts. Such views are currently in the minority.

This paper takes a step in defense of such rival views. Certain features of how thick terms and concepts combine evaluation and description are widely taken to support a semantic view of the relationship between thick terms and concepts and evaluation. I'll focus specifically on the idea that the non-evaluative aspects of their meanings underdetermine their extensions. If they do, then what, if not evaluation, could help to determine the extensions of thick terms and concepts? And how, if not by belonging to their meanings, could evaluations drive their extensions? What I'll argue is that the relevant kind of underdetermination of extension can be expected to arise irrespective of whether the relationship between thick terms and concepts and evaluation is semantic or conceptual and can be explained without that supposition. If that is correct, the underdetermination phenomena I'll discuss cannot be used to support a semantic view of the relationship between thick terms and concepts and evaluation.

8.2 Extension and evaluation

Everyone agrees that thick terms and concepts can be used to convey both non-evaluative description and evaluation. Most writers across various other party lines presume that this is because both aspects are built into the meanings of thick terms and concepts. Information that is 'built into' the meaning of a term is to be understood as including the semantic and conceptual entailments of sentences involving those terms. Assigning non-evaluative information of this kind to thick terms and concepts explains why it is semantically permissible to apply thick terms and concepts only to certain non-evaluatively constrained types of things. For instance, *cruel* can apply only to things that have to do with causing suffering.

The view that evaluation is similarly built into the meanings of thick terms and concepts is often simply assumed as common ground in discussions of other issues about thick terms and concepts, such as the issue of "disentanglement" or "separability." This is the issue of whether thick terms and concepts represent some kind of irreducible fusions of evaluation and description or whether their evaluative and non-evaluative aspects are somehow separable.[2] It is, all the same, a controversial view, especially when taken as a fully general view about thick terms and concepts. This class is usually introduced ostensively by listing some paradigmatic examples, such as *cruel*, *courageous*, *generous*, *greedy*, or those at the beginning of this paper. But already the characterization and the scope of the relevant class are matters of dispute.[3] Some philosophers also argue

---

[2] For discussions of this issue, most of which assume the standard view, see for example Foot (1958), McDowell (1981), Williams (1985), Hurley (1989), Blackburn (1992), Gibbard (1992), Dancy (1995), Elstein and Hurka (2009), Kirchin (2010), and Roberts (2011).

[3] For problems with existing characterizations of what thick concepts are, see Scheffler (1987) and, especially, Eklund (2011). Regarding the scope of the class, there are disputes as to whether ethnic slurs or other sorts of pejorative expression should be counted as thick concepts, and even whether such terms as *unchaste* and *cruel* are significantly alike.

that in fact thick terms and concepts, or at least a very wide range of the paradigmatic examples of them, aren't inherently evaluative in meaning but merely are typically used in ways that convey evaluation.[4] Finally, tacit assumptions to the effect that thick terms and concepts form a uniform class are common in the literature, but such assumptions are rarely defended explicitly and by no means obviously true. It cannot therefore be simply assumed that evaluation is in general built into the meanings of thick terms and concepts.

The view that evaluation belongs to the meanings of thick terms and concepts is often supported by claiming that evaluation "drives the extension" of thick terms and concepts.[5] This former claim may be stated as follows:

Extension The extensions of thick terms and concepts (relative to context) are determined in part by global evaluations.

The extension of a term can be understood as the set of all and only the things that satisfy the term.[6] The qualification "relative to context" will be discussed shortly. The phrase "global evaluations" also requires explanation, regarding both what makes some information evaluative to begin with and what the restriction "global" means.

It is controversial what evaluation is. I'll assume that evaluation is somehow positive or negative in flavor. More precisely, I'll understand evaluation as information to the effect that something has (or lacks) merit, worth, or significance (that is, a positive or a negative standing) relative to a certain kind of standard, namely one that grounds claims of merit, worth, or significance.[7] It is, in brief, information that something is good, or bad, in some way. Being intrinsically good and being instrumentally good, being morally good and being aesthetically good, being good for having features F, G, H, and being a good instance of a kind K, are all ways of being good in the relevant sense. This characterization focuses my discussion on the relationship of thick terms and concepts to evaluations of their instances as good, or bad, in some way.

The restriction to "global" evaluations is meant to focus my discussion on the issue of whether the meaning of *courageous*, for instance, builds in the condition that accepting certain risks of harm for the sake of good goals is good in a certain way, not whether it entails that things falling under it involve accepting certain risks of harm for the sake of

---

[4] See Hare (1952), chapter 7, and (1981), 17–18, 73–5, Blackburn (1992) and Väyrynen (2012).

[5] See for example Foot (1958), McDowell (1981), Williams (1985), Blackburn (1992), Gibbard (1992), Elstein and Hurka (2009), Kirchin (2010), and Roberts (2011).

[6] There is a wrinkle here. It is often not clear in the literature on thick terms and concepts whether extension is understood to be the set of things that actually satisfy the term or the set of the actual and possible things that fall under it. Of these two notions, the latter is more closely aligned with such things as meanings and properties. Thus, insofar as Extension is supposed to bear on the meanings of thick terms and concepts, it may be more charitably understood as concerning this latter notion of extension.

[7] This characterization derives in part from Williams (1985), 125. It takes no stand on whether the evaluative aspects of thick terms and concepts are irreducible to thin evaluations. The mere appearance of the word *good* or *bad* isn't enough to make the evaluation it expresses thin. Consider predicates like *a good philosopher* or *good with children*.

goals whose value is greater than the badness of the harm.[8] The latter evaluation involves evaluative notions that are prior to and independent of *courageous*. Whether φ-ing involves good goals and bad risks can and must be settled prior to settling whether φ-ing is courageous. So the crucial issue whether thick terms and concepts have any *distinctive* significance to evaluative thought and judgment turns specifically on their relationship to global evaluations.[9] The sort of idea that Extension is meant to convey is that the extension of *courageous* is driven in part by some conception (perhaps difficult to articulate) of when and why it is worthwhile to accept a risk of harm for the sake of something valuable. Actions of this general type simply don't count as courageous unless they are thereby good in some way, unless there is some kind of reason to perform them or the like.

Typical arguments for Extension trade on common intuitions about when speakers are engaged in genuine agreement or disagreement about evaluative and normative matters and not merely talking past one another.[10] Consider disputes about distributive justice. Thrasymachus holds (or at least could hold) that distributive justice is conformity to what is in the ruler's interest regarding distributions, while Nozick holds that it is conformity to entitlement and Rawls holds that it is conformity to his two principles of justice.[11] These features of distributions sometimes come apart, so not all three parties can be right about what distributions count as just. But parties to such disputes typically take themselves to be addressing a common topic. Their disagreement about the extension of *distributively just* seems univocal. This cannot be explained by whatever overlap there may be in their conceptions of distributive justice. A better candidate for what ties their dispute together might rather seem to be the assumption that just distributions are distributions with those features, whatever they may be, which make distributions good in a certain way. Their disagreement concerns what the relevantly good-making features of distributions in fact are. This is to treat evaluation as driving the extension of *distributively just*.

I'll briefly register some general concerns about this kind of argument before focusing on other issues. The concern can be illustrated by recalling that Thrasymachus is contemptuous of justice (which he, again, conceives of as what is in the ruler's interest),

---

[8] These evaluative conditions illustrate the distinction between what Daniel Elstein and Thomas Hurka call "global" and "embedded" evaluations. See Elstein and Hurka (2009).

[9] The exception to this is the view that thick terms and concepts are evaluative in some *sui generis* sense in which evaluation need be neither positive nor negative in flavor. I'll set these views aside here; but see Kirchin (this volume) and Roberts (this volume).

[10] A different argument for Extension can be found in Dancy (1995). He infers Extension from his argument that the evaluative and non-evaluative aspects of thick terms and concepts are irreducibly inseparable. The discussion to follow is relevant to this argument as well, since the argument operates with the background presumption that the meanings of thick terms and concepts involve evaluation.

[11] See Plato, *Republic*, Book I, Rawls (1971), and Nozick (1974). I am of course taking liberties in representing Thrasymachus' view as an account of specifically distributive justice. More generally, whether the view described in the text provides the best textual interpretation of Thrasymachus' position in the *Republic* is irrelevant to the illustrative use to which I put the example here. All that is required is that the view is coherent.

but he clearly means to be talking about the same subject as those who praise justice as a virtue and reject his account of it. In other words, it seems that there can be genuine disagreement about what distributions count as just among people who disagree as to whether just distributions are thereby good in a certain sort of way. If that is right, then it cannot be a conceptual constraint that the specific property of distributions, whatever it may be, which is the stuff of distributive justice must be relevantly good-making. As Nicholas Sturgeon notes:

We have . . . many cases in which it is enormously plausible to regard a disagreement as genuine, as focused on a single topic, but in which the dispute is anchored *neither* in a shared set of basic standards *nor* in a shared disposition to (for example) praise justice and condemn injustice.[12]

One general concern about arguments from the conditions of genuine disagreement to Extension is therefore that we often appear to take sameness of topic to be preserved across a broader range of disagreements than these arguments appear to allow.[13]

Another general concern is that what counts as genuine disagreement is itself a controversial topic. It may be unduly narrow to think of disagreement over an utterance as concerning the truth or correctness of its content. Utterances carry many kinds of information beyond the content of the sentence uttered, such as various presuppositions and implicatures. Each of these could potentially underlie a dispute between speakers.[14] The intuition that the speakers are disagreeing can therefore often be explained in a way that doesn't constrain the meanings of the expressions involved or impact their extensions. But in that case it is unclear to what extent intuitions about when speakers are engaged in genuine (dis)agreement about normative or evaluative matters are reliable guides to the meanings or extensions of the terms in which the (dis)agreement is framed. This is but a special case of the general phenomenon that our intuitions about what is said by our utterances may be based on more than what is linguistically determined. The difference between information that is linguistically encoded in the meaning of an expression and information that speakers associate in other ways with its utterances may not always be psychologically salient. Judgments from ordinary speakers may thus fail to match with the distinctions that matter to assignments of meaning or extension.

These general observations generate a further concern about arguments from disagreement to Extension. Intuitions of disagreement are guided in part by how individuals apply and withhold terms. What we have just noted is that patterns of application and withholding may be sensitive not only to the meaning of a term and one's view of its extension but also to various non-semantic factors. If one withholds the application of a term to something, this can be either because one thinks it doesn't fall under the term or because one recognizes that although it falls under the term, applying the term to it

---

[12] Sturgeon (1991), 22; cf. Blackburn (1991), 4–5.

[13] Another general issue here is just to what extent genuine disagreement over a topic requires shared content among claims about the topic.

[14] For discussion, see Sundell (2011).

would be misleading or conversationally inappropriate in some way.[15] If so, patterns of application and withholding for a term may diverge from its extension, including even the extension intended in a stable and cohesive community of users. Thus, even if the application and withholding of thick terms and concepts is sensitive to evaluation, this may well be explicable without treating evaluation as a semantic feature of thick terms and concepts.

I'll bracket these general concerns about Extension for now. Let's instead ask how Extension is supposed to be related to the view that evaluation belongs to the meanings of thick terms and concepts. For the purposes of my discussion this view can be formulated as follows (with elaboration to follow shortly):

Semantic View The semantic meanings of thick terms and concepts involve global evaluations.

While Semantic View entails Extension, the converse isn't true. Linking Extension to claims about the meanings of thick terms and concepts requires further assumptions about how extension relates to meaning.

The nature of meaning is a controversial topic in its own right. But one uncontroversial point is that one cannot use extension to guide views on the identity of concepts: two thinkers can have the same concept but different views concerning its extension, as in our example of distributive justice. Thus, the more closely the meanings of linguistic expressions are related to concepts, the more careful one should be with using extension to guide views about the meaning of an expression or concept identity.[16] Thick terms and concepts are no exception. All that I'll assume about meaning itself is the fairly standard idea that the meaning of an expression or a sentence imposes a set of constraints on what any literal use of it expresses in all normal contexts.[17]

Care is all the more due insofar as meaning may underdetermine extension. Sometimes input from context is needed as well. The meanings of such context-sensitive expressions as indexicals, for instance, don't alone determine their referents on particular occasions of use, since they have different referents in different contexts. The aspect of meaning that remains constant across the different occurrences of context-sensitive expressions in normal contexts is their 'character'—a function from contexts to contents.[18] A general mark of context-sensitive expressions is that their characters

---

[15] For instance, global evaluations might function in typical conversational contexts as background assumptions. If so, they might influence patterns of application and withholding by making only those things conversationally salient which conversational participants treat as admissible candidates for satisfying the relevant evaluation.

[16] The relationship between thick terms and concepts is usually unclear in the literature. A common tacit assumption seems to be that the meanings of thick terms are more or less the same as the concepts they express. (Perhaps the assumption is that terms *have* meanings whereas concepts *are* meanings.) This may be a significant simplification; perhaps, for instance, linguistic meaning is less fine-grained than concepts, in which case the relationship might well be one-to-many rather than one-to-one. See also §§8.4–8.5 this volume.

[17] See for example King and Stanley (2005), as well as Soames (2008), who thinks more strongly that such constraints exhaust semantic meaning. My purposes don't require this latter assumption.

[18] For the notion of character, see Kaplan (1989).

deliver different contents in different contexts. Their extensions vary relative to context.

This suggests the following general view about the connection between the literal meanings of expressions and their extensions:

Meaning-Extension Link Meaning determines extension (relative to context).

I'll take up the issue whether and how thick terms are context-sensitive in sections 8.4–8.5. (Alert readers will notice that I am speaking only of thick terms, and not of thick concepts, as context-sensitive. That is because the label doesn't apply unproblematically to concepts.)[19] But first I want to ask how Meaning-Extension Link is supposed to apply to thick terms and concepts.

The case of Thrasymachus, Rawls, and Nozick suggests that people whom we have no reason to regard as linguistically or conceptually defective can each be talking about distributive justice while failing to share a basic set of standards about what counts as just in distributions of benefits and burdens. Presumably one can then know what *distributively just* means without knowing what things are distributively just, just as one can know what *now* means without knowing to what time it refers in a given context. (Similarly, one can know what *morally good* means without knowing which things are morally good.) How then does meaning determine the extensions of thick terms and concepts?

The question seems to get an elegant solution from Semantic View if we assume Extension. Glossed in terms of the notion of the character of an expression, Semantic View says that the functions from context to content that determine the extensions of thick terms and concepts relative to context involve global evaluations. If Extension is true, then the extension of *distributively just* (relative to context) is most plausibly (given the possibility of genuine disagreement about its extension) taken as driven by substantive evaluative facts about which specific features of distributions are good-making. But Semantic View doesn't require those facts to be involved by having them be built into the meaning of the term. What needs to be built into the meaning of *D is distributively just* is only that *D* has those features, whatever they may be, which make distributions good in the relevant way.[20] This is one way for the meaning of *distributively just* to "involve" the sorts of global evaluations facts which, if Extension is true, are most

---

[19] There are two options regarding how to talk about concepts corresponding to context-sensitive expressions. One option is that if a term like *tall* is context-sensitive with respect to what height counts as tall, then there is no such thing as *the* concept of tallness. There is only the contextually salient concept of tallness, selected from a wide range of more specific concepts given the semantic content of *tall* relative to context. The other option is that there is such a thing as the concept of tallness but it is a Kaplanian character. In the case of *tall*, this function returns no content independently of context and returns different contents in different contexts, thereby helping to specify various specific concepts. This is the sense in which one might talk of "the concept I" if one were so inclined in the case of the first-person pronoun. (Characters resemble concepts in that each is usually located closer to sense than reference.) My present purposes don't require a choice between these options.

[20] See the analysis of *distributively just* in Elstein and Hurka (2009), 521.

plausibly taken to determine its extension. But it seems to explain how one can know what *distributively just* means without knowing which features of distributions in fact are relevantly good-making, and thus how genuine disagreement about its extension, based on disputes about what these features are, is possible.[21] If only one knew which features of distributions in fact are relevantly good-making, one could then figure out which distributions in fact count as just.

This argument for Semantic View raises two issues. Is Semantic View really part of the best explanation of how the extensions of thick terms and concepts are determined? Is Extension itself true, or might the considerations used to support it have some other explanation? I'll focus on a further assumption behind the argument which bears on both issues.

8.3 Underdetermination and evaluation

The best-explanation argument from Extension to Semantic View outlined above involves a significant presumption. It follows from Extension that the non-evaluative (and embedded evaluative) aspects of the meanings of thick terms and concepts under-determine their extensions. But if meaning determines extension (relative to context), then what other further factor but global evaluations could be involved in determining the extensions of thick terms and concepts?

The underdetermination claim that can thus be used to motivate Extension may be stated as follows:

Underdetermination Even the strongest non-evaluative descriptions and embedded evaluations that are built into the meanings of thick terms and concepts underdetermine their extensions (relative to context).

If Underdetermination were false, then there would be no reason to accept Extension, and hence the argument for Semantic View as part of the best explanation of Extension would fail to get off the ground. If, however, Underdetermination is true, that isn't yet enough to motivate that argument for Semantic View. Underdetermination would also be true either if extension (relative to context) were in general underdetermined by meaning or if the meanings of thick terms and concepts *in toto* underdetermined their extensions (relative to context) irrespective of whether their meanings build in evaluation. The presumption behind the best-explanation argument from Extension to Semantic View is therefore that Underdetermination is true specifically in virtue of how global evaluations are involved in their meaning and determination of extension. I'll argue that this presumption isn't needed to explain Underdetermination.

This presumption depends in part on Meaning-Extension Link. Meaning-Extension Link allows that sentences may underdetermine propositional interpretation without

[21] This brackets the earlier general concern about arguments for Extension that are based on univocality intuitions concerning disagreements about extensions.

contextual supplementation when they involve context-sensitive expressions. Context-sensitive expressions include indexicals and demonstratives but may range far more widely, from quantifiers to gradable adjectives to certain verbs.[22] But the presumption is false if linguistic meaning underdetermines extension even relative to context in certain other ways. A particularly stark challenge to Meaning-Extension Link comes from the views of those linguists and philosophers of language who think that the meanings of linguistic expressions systematically underdetermine the contents of utterances, far beyond the gap present in ambiguity, polysemy, and context-sensitivity. Perhaps the meanings of expressions are contextually adjusted to conversational needs through various pragmatic processes that endow expressions with contextual senses, and therefore extensions, which are distinct from their literal meanings.[23] Or perhaps most linguistic expressions have no stable or determinate meanings and determining what counts as an instance of a predicate is in large part a matter of exercising general-purpose abilities that aren't specifically linguistic, and a matter that is systematically up for debate and negotiation in particular conversational situations when different potential assignments of meaning are available.[24] These views imply that extension is in general underdetermined, perhaps radically, by the literal meanings of expressions.

In what follows I'll bracket these kinds of general issues to focus specifically on thick terms and concepts. The assignment of semantic values to sentences involving thick terms and concepts isn't insulated from various more innocent forms of underdetermination. It can be widely agreed that meaning often fails to determine definite semantic value without some help from context, and similar phenomena arise in the case of concepts. If whatever individuates a concept specifies or otherwise determines its content, then the information encoded in its content will typically underdetermine just what entities fall under the concept.[25] Many people have a concept MOUNTAIN that underdetermines just which landforms fall under it (this isn't to say that a determinate technical concept isn't possible!), a concept VEHICLE that doesn't settle whether a Jeep that constitutes a war-memorial statue

---

[22] It matters to Meaning-Extension Link that many of these further cases are controversial. One example is nominal restriction. Modifiers can receive different interpretations depending on what noun they modify. For instance, *huge* can involve a claim about large physical size (*a huge tree*) or the holding of the nominal predicate to a high degree (the more natural reading of *a huge Époisses enthusiast*). This is also true of many uses of *good* and *bad*. The implications of nominal restriction for Meaning-Extension Link depend on how meaning interacts with interpretation relative to context. According to typical forms of semantic contextualism, for instance, nominal restriction engages a contextual parameter that is part of the meaning of the adjective. But according to minimal semantics and many forms of truth-conditional pragmatics, the difference in interpretation is a result of pragmatic processes like enrichment which are underdetermined by meaning. (See, respectively: King and Stanley (2005); Bach (2001) and Cappelen and Lepore (2004); and Carston (2002) and Recanati (2004) and (2010).) These latter views aren't hospitable to Meaning-Extension Link.

[23] See for example the truth-conditional pragmatics of Carston (2002) and Recanati (2004) and (2010) and the lexical pragmatics of Blutner (1998).

[24] See for example the dynamic lexicon view of Ludlow (2008); cf. Larson and Ludlow (1993), von Fintel and Gillies (2011) and Rayo (forthcoming).

[25] This may or may not be due to whatever corresponds to context-sensitivity in the case of concepts.

falls under the rule *No vehicles in the park*, and a concept ATHLETE that doesn't specify whether only humans or also non-human animals engaged in sport fall under it.[26] Which way our concepts should go might be up for debate, not something determined by the concepts themselves.

To see how this goes in the case of thick terms and concepts, consider that a sentence like *x is selfish* entails, as much as a matter of meaning as anything does, at least that *x* manifests some kind of preference for the agent's own happiness or other interests ("happiness," for short) over a greater contribution to the happiness or interests of others. What its meaning, whether evaluative or not, doesn't seem to specify is just how much greater the happiness for others must be relative to the happiness for the agent in order for *x* to count as selfish. Although most people don't think that (other things being equal) spending £100 on a new pair of shoes instead of donating the money to famine aid is selfish, although the latter would benefit others more, it seems that they can fully coherently ask themselves, in a moment of doubt before clicking "Buy now," whether it really is selfish after all. The non-evaluative descriptions that are aspects of the meaning of *selfish* seem only to restrict its application to things with non-evaluative features of a certain generic type, namely bringing about one's own happiness over a greater happiness for others when the latter meets some unspecified differential threshold θ, not to determine any specific property of this type. The question I am addressing is what follows from this sort of phenomenon.[27]

What I'll argue is that insofar as Underdetermination is true of thick terms and concepts, this is an instance of a kind of underdetermination of their extensions (relative to context) by their meanings *in toto* which is exhibited by certain kinds of non-evaluative context-sensitive terms and can therefore be expected to arise irrespective of whether Semantic View is true. I'll also argue that, given the range of contextual factors that are eligible to enter into determining the extensions of thick terms and concepts, Extension is unlikely to hold across all contexts. My conclusion will therefore be that if there is a good argument for Extension or Semantic View as a general thesis about thick terms and concepts, it won't come from Underdetermination.

8.4 Underdetermination and gradability

I'll begin by arguing that the paradigmatic examples of thick terms and concepts express gradable notions and explaining how the kind of context-sensitivity that is characteristic of gradable expressions bears on Underdetermination, Extension, and Semantic View.

Many thick terms and concepts are gradable: they express qualities of which things can have more or less and can thus be used to order the things under discussion.

---

[26] The last two examples are due to Hart (1958) and Ludlow (2008), respectively. They show that this kind of underdetermination doesn't reduce to vagueness.

[27] The phenomenon itself is well discussed by Elstein and Hurka (2009).

Adjectives are one standard way to express gradable notions. Many thick terms are adjectives that exhibit syntactic marks of gradability. They admit of comparatives: things can be *more frugal* or *more courageous* than others, *less cruel* or *less generous* than others. And they take degree modifiers: things can be *very* industrious, *somewhat* tacky, *extremely* frugal, and so on. But gradability isn't a property of adjectives alone. One person can be *more of* a jerk than another, or a *bigger* smoker than another, one can regret some things *more than* others or regret them *very much*, just as a person can be an *outright* idiot or an *absolute* genius and an attempt at a clever joke can be a *total* failure.[28] And expressions that aren't syntactically gradable might still be semantically linked to scales of measurement in the way gradable expressions are. Insofar as thick terms and concepts are gradable, some of their features might be explicable as features of gradable expressions in general. The question will be how their gradability bears on Semantic View.

According to the standard treatment of gradable adjectives, their semantic interpretation involves three operations: determining a dimension (the quality of which there can be more or less), computing a scale (an ordering with respect to the dimension) and computing a standard (a value on the scale that is high enough to count).[29] For instance, *tall* orders things according to their (ascending) height. This analysis makes straightforward sense of comparatives: *A is taller than B* can be analyzed as saying that the value *A* takes on the scale of tallness (which is something like *A*'s degree of height) exceeds the value *B* takes on the scale of tallness. The positive form is usually treated as implicitly comparative: *A is tall* can be analyzed as saying that the value *A* takes on the scale of tallness exceeds the contextually determined threshold for counting as tall. The standard will be such that the objects of which the positive is true "stand out" in the context of utterance relative to the relevant measurement.[30]

We can accordingly distinguish at least three issues concerning the interpretation of gradable thick terms like *courageous*, *cruel*, and *selfish*:

---

[28] I am not suggesting gradability as a diagnostic for thick terms and concepts. For instance, *morally good* is a gradable expression but is typically classified as thin. Moreover, some thick concepts may lack gradable expressions; *murder* doesn't seem to be a gradable term but some think that murder is a thick concept. Focusing on gradability works here as a heuristic guide.

[29] Any respect in which two things may be compared may qualify as a standard in the relevant sense. It needn't be evaluative in the sense discussed earlier in §8.2.

[30] This sketch draws primarily on the degree-based analysis in Kennedy (2007) and Glanzberg (2007). Earlier degree-based accounts of gradability include Cresswell (1977), von Stechow (1984), and Bierwisch (1988a), (1988b), and (1989). Such accounts can treat gradable adjectives and their comparatives either as measure *functions* from individuals to values on a scale or as *relations* between individuals and such values. There is also room for debate about whether scale values are to be understood as abstract objects or not, as degrees or intervals, and so on. A degree-based analysis is most naturally understood as requiring that the comparative generate a total ordering (Kennedy 2007). But this seems too strong as a general requirement on scale structure (van Rooij 2011).

The main alternative to a degree-based analysis is to analyze gradable adjectives as simple predicates whose extension varies with respect to a contextually determined comparison class. See Klein (1980); cf. Ludlow (1989). One worry here is that there seems to be no reason why a standard should have to be based on a comparison class. See DeRose (2008).

- *Standards:* What degree of courageousness (etc.) is minimally sufficient to count as satisfying *courageous* (etc.), and hence determines its extension, relative to context?
- *Dimensions:* What counts as courage (etc.)? What is the quality of which there may be more or less, or a lot or little?
- *Scale structures:* What counts as more (or less) courageous (etc.) than what? What determines how two objects are related on the scale?

We can further ask how these issues bear on Extension and Semantic View. The issue of how the meanings of thick terms and concepts are related to dimensions will occupy section 8.5. Issues about scale structure are the least central to my present concerns and will appear mainly in notes. How the meanings of thick terms and concepts are related to standards is the clearest of the three issues, so I'll discuss it first.

Gradable expressions are generally context-sensitive with respect to the standard. For instance, Amy may count as tall in a discussion of ballerinas but not in a discussion of basketball players. So the extension of *tall* varies with context. Many thick terms are context-sensitive in this way. What counts as satisfying *generous*, for instance, is different when millionaires and paupers are in question.[31] Help from context is required to set the standard. There is no such thing as the set of generous things, except relative to context.[32] Thus, insofar as positive constructions involving thick terms vary in extension from one context to another because they make reference to a degree that exceeds a contextually specified standard, this delivers Underdetermination. But that will hold irrespective of whether thick terms are inherently evaluative in meaning, and so explaining it doesn't require Semantic View.

One might nonetheless think that in fact the contextually supplied standards for satisfying thick terms are determined in part by evaluation. This would support at most Extension. It wouldn't support Semantic View because not all literal uses of gradable expressions in normal contexts make reference to a degree that exceeds a contextually specified standard; at most their positive forms do.[33] Such reference is therefore not encoded in the meanings of gradable thick terms. But even the case of Extension is unclear. Context-sensitive expressions differ with respect to whether their meanings specify what contextual inputs are relevant to determining their semantic values. It is therefore an open question whether it is built into the meanings of thick terms that global evaluation plays (or doesn't play) such a role.

---

[31] Some may be inclined to think that if something displays the property T measured by a thick term *T* to any degree, then it is automatically an instance of *T*. They should note that something can have a value on the scale of T-ness without counting as T. Such sentences as *He couldn't muster up enough courage to save his comrades* seem to make perfect sense. Compare *tall*: some degrees of tallness (that is, values on the dimension of ascending height) qualify as short in some contexts.

[32] This makes some writers' focus on what it takes to "master the extension" of a thick concept, in a way that enables ones to apply it correctly to new cases, seem ill-conceived. This focus can be seen in, for example, McDowell (1981), Williams (1985), Dancy (1995), Kirchin (2010), and Roberts (2011).

[33] Each of *Bill is tall*, *Bill is taller than Ted*, and *Bill is as tall as Ted* contains the adjective *tall*, but neither of the latter two constructions (the comparative and the equative) makes reference to a degree that exceeds a contextually specified standard; only the positive construction does so. See Rett (2008).

This complication arises because we must distinguish semantics from "metasemantics". The job of semantics is to specify what kind of meaning each sentence has as some compositional function of the semantic values of the constituents of the sentence.[34] This calls for an assignment of semantic values relative to contexts. The job of metasemantics is to say how constituents wind up having the semantic values they do—how context fixes a particular value for an expression in context. The semantics of context-dependent parameters is simply that they are set to values by context. Such issues about the interpretation of gradable adjectives as how the standard is computed, and on the basis of what sorts of rules and contextual inputs, are metasemantic in this sense. So is, therefore, the issue whether evaluation plays a standard-setting role.

Context-sensitive expressions differ with respect to how direct their metasemantics is. One example of a fairly direct metasemantics are pronouns like *he* and *she*. Their metasemantics involves pronoun resolution rules, which help determine to which individual a pronoun refers in a particular context. These rules might direct us, for instance, to the speaker's referential intentions to determine semantic value. In general, how context manages to set the values of contextual parameters is something that may be constrained, but is typically not fully determined, by the meaning of the expression in question.

The meanings of gradable expressions generally show few restrictions on how the value of the standard parameter may be set, beyond there having to be some appropriately salient factors in the context. For instance, speaker intentions may be neither necessary nor sufficient for setting it, although they often play a role; some speakers' intentions may be too idiosyncratic relative to the presuppositions shared by other speakers or the broader structure or aims of the discourse.[35] Working out the standard value from context may require taking into account a range of factors which may combine in complex ways. Factors that may in general play a role in semantic interpretation include: salient properties of the context; the denotations of the expressions involved; whatever intentions and interests to compare and classify speakers and hearers might have, plus coordinating intentions; the structure of the discourse in which the utterance appears; and a rule that Christopher Kennedy calls the "principle of interpretive economy," which requires making as much use as possible of the (conventional) meanings of expressions in computing what a sentence expresses in context.[36] There is no reason in advance to think that gradable thick terms will differ from other gradable expressions in this respect.

All this matters here because if the meanings of gradable expressions don't generally specify the contextual factors that determine the standard, then nothing about

---

[34] Semantics that is "descriptive" in this sense may be distinguished from accounts of the factors (such as inferential role, causal-historical profile, or whatnot) in virtue of which particular expressions have the semantic properties they do. This is sometimes called "foundational semantics," but sometimes it, too, is called metasemantics. I won't follow this usage.

[35] Glanzberg (2007), 24.   [36] See Kennedy (2007) and Glanzberg (2007).

gradability gives us grounds to think that the meanings of thick terms specify that global evaluations must play a role in setting the standard relative to context. Global evaluations might play no role, or they might play a role sometimes but not always, depending on context. This raises a challenge to the claim that the meanings of thick terms involve global evaluations in their interpretation.

Let's use *selfish* to illustrate how thick terms might not specify as a matter of semantic rule that global evaluations play a standard-setting role. Impartialists in ethics think that preferring n units of happiness for oneself over $n^\star$ units of happiness for others is wrong whenever $n^\star > n$, whereas those who advocate agent-centred prerogatives think that it is wrong only when $n^\star$ exceeds some higher threshold θ above n. They agree that self-ishness is gradable. But, for all that is at issue between them as regards wrongness, they needn't disagree over whether choosing ten units of happiness for oneself over eleven units of happiness for others is selfish.[37] Agent-centrists could agree that it is selfish, but permissibly so. Or, impartialists could agree that it isn't selfish, but judge it to be wrong even so. This would be enough to explain how they can use the term *selfish* univocally despite their disagreement over wrongness, without treating the meaning of *selfish* as necessarily co-opting global evaluations.

What if impartialists and agent-centrists *did* take the moral sticking point to concern which acts are selfish? In that case their disagreement would be about how something must rate on the metric of happiness differential to count as selfish—that is, about the standard for counting as selfish. Would this kind of disagreement require global evaluations to play a role in determining the standard, as a matter of semantic rule or not?

Nothing in the semantics of gradability *rules out* that global evaluations can enter into determining the degree exceeding which counts as selfish in context. If impartialists and agent-centrists want to couch their dispute as concerning which specific property of preferring one's own happiness over a greater happiness for others counts as selfishness relative to context, then they can agree for the purposes of conversation that the stuff of selfishness is, for instance, those preferences for one's own happiness over a greater happiness for others which are wrong *according to the correct moral standards, whatever they are*. This kind of broadly objectivist presumption could well be salient in typical moral contexts.[38]

But equally nothing in the semantics *requires* that global evaluations play a standard-setting role in the interpretation of thick terms.[39] As far as I can tell, it is semantically permissible to take social convention rather than morality to determine what counts as

---

[37] This departs from the stipulation in Elstein and Hurka (2009), 522, to whom I owe the example.

[38] Compare the version of contextualism about deontic modals such as *ought* in Dowell (ms).

[39] It also seems that in certain contexts at least standards concerning such thin notions as wrongness *cannot* play a standard-setting role with respect to *selfish*. Whenever a selfish act must have whichever specific feature of preferring one's own happiness over a greater happiness for others that makes acts wrong, an act cannot be wrong *in virtue of* being selfish. For in that case a selfish act is trivially a wrong act, and it seems that if *A*s are trivially *B*s, then something cannot be a *B* in virtue of being an *A*. Similarly, the judgment that the act is wrong couldn't be justified by saying that it is selfish, if it can only count as selfish to begin with if it is wrong.

selfish in some contexts, or for psychologists studying self-regarding and other-regarding behaviors to set an operational standard of selfishness without being guided by global evaluations. All that the semantics requires is that there be some degree of concern for one's own happiness over a greater happiness for others which exceeds a contextually determined standard. The rest depends on matters that aren't in general settled by the meanings of thick terms. So nothing here implies that the meanings of gradable thick terms in general involve global evaluations in determining the contextually specified standard; nothing here supports even Extension as a general thesis about thick terms and concepts.

So far I have argued that many thick terms express gradable notions, and that because such expressions are context-sensitive, their extensions are underdetermined by their meanings *in toto*, irrespective of whether their relationship to evaluation is semantic. I also argued that nothing in the context-sensitivity of gradable expressions shows that global evaluations are required in determining the contextually supplied standards for counting as satisfying thick terms. We have yet to see anything about Underdetermination that would support either Extension or Semantic View as a general thesis about thick terms and concepts.

At this point it is natural to wonder why the sort of underdetermination that would best support Semantic View or even Extension should concern the contextually specified standard. For instance, although Nozick and Rawls can agree that distributive justice is gradable, their disagreement isn't about how just a distribution has to be to count as just. Rather it is about what feature of distributions is the stuff of distributive justice. This points to a different possible form of underdetermination having to do with the dimensions associated with thick terms and concepts:

Underdetermination + The non-evaluative aspects and embedded evaluative aspects of the meanings of thick terms and concepts underdetermine the metric on which such a term grades things (relative to context), not just the value on the metric needed to make the grade which determines the term's extension (relative to context).[40]

Underdetermination + implies that the non-evaluative and embedded evaluative aspects of the meaning of *courageous*, for instance, underdetermine what counts as courage—the quality of which there may be more or less, (not) enough or too much, very much and so on, and different amounts of which may be enough in different contexts to count as courageous. This couldn't be explained in terms of gradability, since gradability only requires that the (more or less complex) qualities or properties ascribed by thick terms may be realized to different degrees. Thus, insofar as Underdetermination + is what is typically going on in cases where evaluation might be regarded as driving the extensions of thick terms, that cannot be explained by gradability as such either.[41] This seems right.

---

[40] I use the word "metric" to refer to dimensions to capture the phenomenon of multidimensionality that I'll discuss shortly.

[41] Thanks to Daniel Elstein for emphasizing the points in this paragraph.

Not all thick terms and concepts exemplify Underdetermination + with equal plausibility, however. Underdetermination + isn't plausibly true of *selfish*, for instance. The semantic clause that selfish acts somehow involve preference for the agent's own happiness or interests over a greater contribution to other people's happiness or interests (where "somehow" signals that the happiness differential threshold θ is left semantically unspecified) underdetermines the extension of *selfish*. But that clause is enough to generate a multitude of such comparative facts as that preferring ten units of happiness for oneself over twenty units of happiness to others is more selfish than preferring it over eleven units of happiness to others and as selfish as preferring five units of happiness for oneself over fifteen units of happiness for others. Such comparative and equative facts are enough to generate a scale of selfishness. Since a scale is an ordering along a dimension, those facts are enough to determine a metric of selfishness without global evaluation. So Underdetermination + isn't plausible with respect to *selfish*. It is at most plausible with respect to some restricted set of thick terms and concepts. I'll now turn to discuss what Underdetermination + shows in cases where it seems plausible.

8.5 Explaining Underdetermination +

Underdetermination + makes a claim about how the metrics along which thick terms and concepts measure things are determined. To assess this we need to know how the properties ascribed by gradable expressions are in general determined. Semantic theory doesn't much care about how or why different gradable expressions differ in the ways they do.[42] The standard analysis of gradable adjectives takes no stand on what properties are scaled by a particular scale, what contextual factors qualify as inputs for computing the scale, and what computational rules must be taken into account and how these must be weighed in fixing the values of such contextual parameters as the standard. All it requires is that there *be* a scale and standard somehow to be computed from context; determining these is, again, the job of metasemantics.

The fact that *tall*, *frugal*, *cruel*, and *good* are used to measure different qualities is due to differences in their conventional meanings. (This point reflects the principle of interpretive economy mentioned earlier.) How fully their meanings determine the associated dimensions varies from case to case. In some cases the dimension is conventionally fixed; examples include *tall* (ascending height), *young* (descending age), and *cheap* (descending cost). The same applies to comparisons and the corresponding scale structure: whether one thing is taller, heavier, younger, or cheaper than another is usually clear. In other cases, however, meaning underdetermines the metric. For instance, *good* and *bad* require help from context to determine a respect of comparison, such as being good at cooking or with children, or what is a good government policy for higher education.

---

[42] Main exceptions concern certain structural features of scale structure, such as the distinction between "absolute" and "relative" gradable adjectives; see Kennedy (2007).

Even fixing a respect of comparison may, however, underdetermine the relevant property. For instance, the qualities measured by *a good philosopher*—such as perhaps insight, creativity, clarity, and rigor, among others—may themselves be unclear or controversial. It may similarly be unclear or controversial how the corresponding scale is structured, due to the possibility of disputes about who is a better footballer or a cook than who, which are based on disagreements about what the qualities to be measured are or their relative importance.

These complications arise clearly in the case of "multidimensional" notions, such as what is a good philosopher. Many gradable expressions grade things along multiple dimensions in one and the same context and the degree to which things possess the quality measured by the expression depends somehow, on the basis of some kind of combinatorial function, on separate orderings along these multiple dimensions. Getting a metric along which things are graded requires such a function.[43] For instance, whether somebody is more philosophically talented than somebody else depends, somehow, on separate orderings along such dimensions as insight, creativity, clarity, rigor, and more, and how painful something is depends, somehow, on the intensity and the duration of pain. Just how the various dimensions are to be combined is often not clear. So multidimensionality tends to complicate the determination of both scale structures and the metrics along which things are graded.

I'll now argue that the meanings of multidimensional thick terms and concepts *in toto* tend to underdetermine their extensions relative to context. The argument will apply even to their comparatives, because one factor in play will turn out to be variation in combinatorial functions. This is important because the comparatives of unidimensional gradables are typically not context-sensitive. The extension of *taller than* doesn't vary with context. The upshot will be that if Underdetermination + is true of thick terms and concepts, this will be so for reasons that hold irrespective of whether the relationship between thick terms and concepts and evaluation is semantic. Thus Underdetermination + fails to support Semantic View. As before, global evaluations may play a role in determining the extensions of thick terms and concepts, but there is no reason why they need to. Underdetermination + therefore fails to support Extension as a general thesis.

As with Underdetermination, there are various general views about word meaning that could be used to make the kind of point I am seeking to make about Underdetermination +. For instance, various people argue that word meanings are systematically underdetermined by what is common coin among speakers and protean beyond certain fairly minimal constraints.[44] This fits with the observation that the shared meaning of thick terms and concepts tends to consist in various hints and clues. Witness such typical lexical

---

[43] The relevant metrics, and locations in them, can be represented using n-dimensional vectors, where the combinatorial function is a mapping from such vectors to degrees, positions, or values on a scale; see Benbaji (2009), 321–3. Formulating such vectors can be a real challenge.

[44] See for example Blutner (1998), Carston (2002), Ludlow (2008), von Fintel and Gillies (2011), and Rayo (forthcoming).

entries as *disposed to inflict suffering, indifferent to or taking pleasure in another's pain or distress* in the case of *cruel*; *lacking consideration for others, concerned chiefly with one's own personal profit or pleasure* in the case of *selfish*; *having or showing a tender and considerate and helpful nature* in the case of *kind*, and *not deterred by danger or pain; strength in the face of pain or grief* in the case of *courageous*.[45] It might be thought to be up to conversational participants to deploy further cognitive resources to flesh out these largely non-evaluative clues in some way that is appropriate to their discourse situation. This might involve selecting meanings from a pre-existing stock of different possible senses or generating them anew.

Thick terms operate primarily in domains which we regard as normatively significant. Ascriptions of justice and cruelty, for instance, aren't usually neutral with respect to what is good or bad, but are at least normally taken to ascribe good-making and bad-making features that are connected to reasons for action. It would then be no surprise if conversational partners often relied on global evaluations to flesh out contextual meanings for thick terms even if the relevant global evaluations weren't contained in their meanings. Insofar as our evaluative outlooks tend to overlap, it would also be no surprise if evaluation commonly guided us to certain particular sharpenings among the many available. There would, however, be nothing privileged about those sharpenings.[46] They would come about through various general norms that regulate how word meanings are negotiated and sharpened against conversational purposes and context, through having the views of particular speakers expressed and debated, and on that basis accepted, rejected, or modified.[47]

Nothing in this shows more than that thick terms can be loaded with evaluation in particular discourse situations. There is no reason why fleshing out a meaning that works for a discourse situation would have to involve global evaluation. In some discourse situations the extension of a thick term may be driven by global evaluation, but in others global evaluation will have to be trimmed off for conversational coherence and progress to be possible. An example would be a discussion of how we should respond to international terrorism in which someone claims that we should respond to it with focused brutality.[48] Such a speaker may be misguided but seemingly doesn't have to be conceptually confused. It won't do to interpret *brutal* as building in negative global evaluation in such a discourse situation. Nothing in this kind of general picture portrays evaluation as something contained in the meanings of thick terms.

My purposes don't, however, require any such general view about word meaning. The example concerning approval of focused brutality, for instance, can stand on its

---

[45] See for example *The Oxford English Dictionary* or WordNet 3.0 (a lexical database for English accessible at <http://wordnet.princeton.edu/>, accessed 1 April 2012).

[46] There would, accordingly, be no privileged concept CRUEL, SELFISH, or COURAGEOUS that could somehow override various factors that bear on deciding how to sharpen the corresponding terms relative to conversational goals, interests, and context.

[47] See for example Ludlow (2008), 125–6. Also metasemantic decisions as to what contextual factors to take into account in semantic interpretation are often subject to debate and negotiation. See Glanzberg (2007) and Sundell (2011).

[48] This example was relayed to me by Nicholas Sturgeon many years ago. A real-life example reported to me by Remy Debes speaks to the same point, namely the comment, "Don't change this beautiful and brutal sport" made in a discussion of whether increased concussion risks in (American) football should lead to changes in the rules of the game.

own. Moreover, I'll now argue that the constraints on literal uses of multidimensional terms in normal contexts tend to underdetermine the metrics along which things are graded, and thereby their extensions relative to context. This will support Underdetermination + in the case of multidimensional thick terms and concepts irrespective of whether their relationship to global evaluation is semantic.

How bald someone is depends on both the number of hairs on his scalp and the distribution of hairs on it. How painful something is depends on both the intensity and the duration of pain. The extent to which something is bald or painful is some combinatorial function of separate orderings on the relevant dimensions. But the relative weighting of these dimensions, and hence the resulting combinatorial function, can vary with context. The shape of this function will be constrained by whatever formal properties the resulting ordering of objects is desired to have.

Whatever shape such functions take, it seems rare for them to be specified as a matter of meaning or even linguistic practice. How the various dimensions are to be weighted against one another often depends on context. Although a person with a greater number of hairs on his scalp usually counts as less bald than a person with fewer hairs on his scalp, this isn't so when the former's hairs are distributed on his head very unevenly in one big tuft.[49] Or consider *painful*. Day 1 has a longer duration of pain of lower intensity; Day 2 has short durations of pain of higher intensity. Thus Day 1 ranks higher in the duration of pain but Day 2 ranks higher in the intensity of pain. Which is more painful of the two? The answer may vary with context even if Day 1 is stipulated to have a greater total amount of pain, let alone if the total amounts are stipulated to be equal. The extension of *more painful than* may vary with context in this way because the dimensions of intensity and duration may be weighted differently in different contexts. And the question of which is more painful appears to have no answer unless some idea of how intensity and duration are to be weighted against one another is supplied. If so, then not only the positive construction *D is a painful day* but also the comparative construction *D1 is a more painful day than D2* may vary in extension relative to context. It is exceedingly difficult to see facts about which dimensions we take into account in which contexts as determined by meaning or linguistic practice.[50] I conclude that the meanings of multidimensional gradable expressions *in toto* only require that there *be* a metric along which things are graded and tend to underdetermine the metric.

Thick terms seem no different. What counts as courageous, for instance, is some function of the likelihood and (relative) value of the goods to be achieved by action,

[49] See Wasserman (2004), 396.

[50] The points in this paragraph are further reinforced by the observation that in some contexts one of the dimensions may even drop out of the comparison as irrelevant to the conversational purpose. We often don't focus on all dimensions when evaluating comparatives, and often focus just on dimensions along which things differ and ignore those along which they don't. See for example Benbaji (2009) and van Rooij (2011).

the likelihood and (relative) disvalue of the harms risked by action, confidence, and resistance to fear and so on. Competent speakers know that the meaning of *courageous* permits greater feelings of fear the greater the danger, at least provided that the expected gains of going ahead are greater than its expected harms. They also know that absolute fearlessness or confidence in the face of a high probability of a grave injury in pursuit of something barely worthwhile is reckless or foolhardy, not courageous. And they know that fearfulness in the face of a low probability of grave injury in pursuit of something highly worthwhile is cowardly, not courageous.[51] But even if these are stable constraints on what literal uses of *courageous* express in normal contexts, it is exceedingly hard to believe that meaning or linguistic practice specify much further which dimensions are to be taken into account and how the relevant dimensions are to be weighted in a particular context, or otherwise specify how changes along each contextually relevant dimension are related to changes of location in the multidimensional property space out of which the relevant metric is constructed. They seem not to determine the relevant combinatorial function much beyond requiring that there be one and placing some loose constraints on it. Surely *courageous* is also not an isolated case.

What does all this mean for the relationship between thick terms and concepts and evaluation? It is consistent with this argument that some thick terms and concepts contribute evaluative properties as dimensions and even that some of them do so as a matter of meaning. I see no a priori argument to the contrary. But as with the case of standards, nothing in the semantics of thick terms in general requires global evaluations to play a role in specifying the metrics along which multidimensional thick terms grade things. Factors that determine them seem not to be stable across the literal uses of thick terms in normal contexts in the way they should be if they were built into meaning. Whether, and how widely, global evaluations play a role in determining those metrics are primarily

---

[51] Many issues concerning scale structure arise here. One is whether thick terms impose scales that are continuous with the scales imposed by their antonyms, where they have one. It is by no means clear how such clusters as *brave*, *bold,* and *courageous* relate not only to *cowardly*, *timid,* and *fearful* but also to *reckless*, *foolhardy,* and *rash*. How to analyze those virtue terms that have two different kinds of vice terms as their negative counterparts is an important issue. Depending on how this plays out, *courageous* may not be semantically well represented as denoting a mean between cowardice and recklessness in the sense that it maps its arguments onto intervals on a scale which lie between intervals to which *cowardly* and *foolhardy* map their arguments. We might not want to analyze the relevant metric so that foolhardy things rank higher than courageous things on it. Foolhardy things may display greater confidence in the face of danger and fear than courageous things, but this isn't the same as having a greater degree of courage than courageous things. If *courageous*, *cowardly*, and *foolhardy* don't impose a scale on the same dimension, then they are to be represented as operating on different (albeit related) scales. If they do, then the issue will be how to represent the scales imposed by *courageous*, *cowardly*, and *foolhardy* as scales on the same dimension. It doesn't seem that either cowardly or foolhardy things have in any systematic way more or less of the same quality than the other. Irrespective of how *cowardly* and *foolhardy* compare to one another, it also isn't clear how to represent them relative to the standard for counting as courageous. A further issue will be how to analyze such sentences as *A is as cowardly as B is foolhardy* in terms of degrees on the same dimension. All these issues deserve more attention.

questions about contexts and metasemantics. Since contexts are contingent entities, there seems to be no reason to think in advance that global evaluations must always be a relevant factor. Because all these considerations arise from reflection on multidimensional gradable terms in general, they imply that Underdetermination + has nothing in particular to do with how thick terms and concepts combine description and evaluation. The exception would be cases where global evaluations are built into the meanings of thick terms and concepts. Whether isolated cases of this kind would be able to carry any significant broader metaethical consequences is a complicated issue which I cannot address here. But since exceptions are sometimes best left as spoils to the victor, the question merits attention.

I conclude that insofar as Underdetermination + holds for thick terms and concepts, that is generally because the meanings of multidimensional thick terms *in toto* underdetermine the metrics along which things are graded. This account is simpler and based on more general principles than the explanation that the meanings of thick terms fully determine those metrics by co-opting global evaluations that complement the non-evaluative aspects of their meanings in determining them. The explanation I have developed works irrespective of whether the relationship between thick terms and concepts and evaluation is semantic. Therefore Underdetermination + doesn't support Semantic View as a general thesis about thick terms and concepts. It doesn't even support Extension. For although global evaluations may play a role in determining the extensions of thick terms relative to context, nothing in the semantics forces this.[52]

This explanation of Underdetermination + has broader dialectical significance. Consider, for instance, cases where global evaluation seems to drive the extensions of thick terms and concepts. If Underdetermination + can be explained without appealing to Extension, then attempts to explain these cases as cases where patterns of application and withholding are influenced by non-semantic or conversational factors cannot be dismissed out of hand.[53]

The sort of explanation I have developed also undercuts certain arguments regarding the putative implications of the underdetermination phenomena discussed in this paper. Allan Gibbard, who explicitly allows that there might be non-evaluative constraints on a thick concept, claims that statements that predicate thick concepts of things don't have "enough of" descriptive meaning that, in some combination with evaluative meaning, can yield the full meaning of the statement. In thick concepts, "descriptive and evaluative components intermesh more tightly than that."[54] Jonathan Dancy similarly claims that the non-evaluative aspects of the meanings of thick terms and concepts determine only the range or domain in which they operate, but no content that could stand alone in semantic evaluation.[55] For instance, describing courage as "something to do with

---

[52] Moreover, as mentioned earlier in n. 39, there seem to be contexts where at least certain kinds of evaluations don't, or even can't, play such a role.

[53] I alluded to the availability of such explanations in §8.2.

[54] Gibbard (1992), 277–8.   [55] Dancy (1995), 275–7.

confidence, resistance to fear and danger," or lewdness as "something to do with sexual display, something to do with mockery" doesn't determine a meaning or content for *courageous* or *lewd*. But all that one needs to do to explain why such general characterizations underdetermine a content that can stand alone in semantic evaluation is to point out that they try to characterize terms that are context-sensitive in abstraction from any particular context. The indeterminacy of characterizations that merely specify the various dimensions shows neither that evaluation belongs to the meanings of thick terms and concepts nor even that it drives their extension. (Dancy uses their indeterminacy explicitly to argue for Extension.) If *courageous*, for instance, is multidimensional, then we shouldn't expect to be able specify in the abstract how its various dimensions are to be combined and weighted in particular cases.

I am therefore underwhelmed by arguments for Extension or Semantic View that are premised on the observation (pressed by both Dancy and Gibbard) that linguistic convention and practice attach no sharp non-evaluative properties to thick terms and concepts. That is generally not the case with context-sensitive terms, especially if they are multidimensional. It is also useful to note that views that deny Extension and Semantic View aren't committed to being able to specify or characterize the properties ascribed by thick terms and concepts in "thickness-free" terms. If the meaning of *selfish* doesn't build in global evaluation, then *selfish* itself ascribes non-evaluative properties in contexts where it takes such properties as its semantic values. But, like a wide range of ordinary terms, both evaluative and not, *selfish* may have no informative analysis in independently intelligible *selfish*-free terms. Similarly, a term like *courageous* might easily lack any accurate *courageous*-free correlate irrespective of whether its relationship to the global evaluations it may be used to convey is semantic.

8.6 Conclusion

I have argued that insofar as the non-evaluative aspects of the meanings of thick terms and concepts underdetermine their extensions, this is due to their meanings *in toto* and can be expected to arise irrespective of whether their relationship to global evaluations is semantic. Considerations of underdetermination fail therefore to support the hypothesis that their meanings in general build in global evaluations, and support the hypothesis that their extensions are driven by global evaluations at most with respect to certain contexts. I wish to close by noting that thick terms and concepts may be quite diverse in their relationship to evaluation. For all I have argued here, some may bear a semantic relationship to global evaluation, but we have yet to see any good reason to make such semantic posits across the class of thick terms. The issue would be an empirical question about the meanings of particular thick terms, in the way questions about semantic posits generally are questions about the meanings of particular linguistic expressions. The construction of the metrics along which things are graded may be guided by evaluation, but it may not, depending in part on whether a thick term is multidimensional and in part on context. The standard for making the grade on the metric may be driven by evaluation, but again

158　THICK CONCEPTS

it may not, depending on context. And so on. Further dimensions of diversity may well come into view in future inquiry.

The lack of a one-size-fits-all account of thick terms and concepts would have important ramifications all by itself. The greater the room for differences in the sorts of respects just enumerated, both across different thick terms and concepts and across different contexts of their use, the harder it is to see why the global evaluations that thick terms and concepts may be used to convey should in general be treated as built into their meanings. The extant literature on thick terms and concepts conspicuously fails to entertain the possibility that diversity among thick terms and concepts goes deep but has no systematic theoretical upshot. This is beginning to look like a serious possibility.[56]

References

Bach, Kent (2001) 'You Don't Say?' *Synthese* 127: 11–31.

Benbaji, Yitzhak (2009) 'Parity, Intransitivity, and a Context-Sensitive Degree Analysis of Gradability', *Australasian Journal of Philosophy* 87: 313–35.

Bierwisch, Manfred (1988a) 'Tools and Explanations of Comparison—Part 1', *Journal of Semantics* 6: 57–93.

—— (1988b) 'Tools and Explanations of Comparison—Part 2', *Journal of Semantics* 6: 101–46.

—— (1989) 'The Semantics of Gradation' in Manfred Bierwisch and Ewald Lang (eds.) *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation* (Berlin and Heidelberg: Springer-Verlag), 71–261.

Blackburn, Simon (1991) 'Just Causes', *Philosophical Studies* 61: 3–17.

—— (1992) 'Morality and Thick Concepts: Through Thick and Thin', *Proceedings of the Aristotelian Society, Supplementary Volume* 66: 285–99.

Blutner, Reinhard (1998) 'Lexical Pragmatics', *Journal of Semantics* 15; 115–62.

Cappelen, Herman and Ernest Lepore (2004) *Insensitive Semantics* (Oxford: Blackwell).

Carston, Robyn (2002) *Thoughts and Utterances* (Oxford: Blackwell).

Cresswell, Max (1977) 'The Semantics of Degree' in Barbara Partee (ed.) *Montague Grammar* (New York: Academic Press), 261–92.

Dancy, Jonathan (1995) 'In Defense of Thick Concepts', in Peter A. French, Theodore E. Uehling, and Howard K. Wettstein (eds.) *Midwest Studies in Philosophy* XX: 263–79.

DeRose, Keith (2008) 'Gradable Adjectives: A Defence of Pluralism', *Australasian Journal of Philosophy* 86: 141–60.

Dowell, Janice (forthcoming) 'Flexible Contextualism about Deontic Modals', *Inquiry*.

Eklund, Matti (2011) 'What Are Thick Concepts?' *Canadian Journal of Philosophy* 41: 25–49.

Elstein, Daniel and Thomas Hurka (2009) 'From Thick to Thin: Two Moral Reduction Plans', *Canadian Journal of Philosophy* 39: 515–35.

Foot, Philippa (1958) 'Moral Arguments', *Mind* 67: 502–13.

Gibbard, Allan (1992) 'Morality and Thick Concepts: Thick Concepts and Warrant for Feelings', *Proceedings of the Aristotelian Society, Supplementary Volume* 66: 267–83.

Glanzberg, Michael (2007) 'Context, Content, and Relativism', *Philosophical Studies* 136: 1–29.

Hare, R. M. (1952) *The Language of Morals* (Oxford: Clarendon Press).

—— (1981) *Moral Thinking* (Oxford: Clarendon Press).

Hart, H. L. A. (1958) 'Positivism and the Separation of Law and Morals', *Harvard Law Review* 71: 593–629.

Hurley, S. L. (1989) *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press).

Kaplan, David (1989) 'Demonstratives' in Joseph Almog, John Perry, and Howard Wettstein (eds.) *Themes from Kaplan* (Oxford: Oxford University Press), 481–563.

Kennedy, Christopher (2007) 'Vagueness and Grammar: The Semantics of Relative and Absolute Adjectives', *Linguistics and Philosophy* 30: 1–45.

King, Jeffrey and Jason Stanley (2005) 'Semantics, Pragmatics, and the Role of Semantic Content', in Zoltán Szabó (ed.) *Semantics Versus Pragmatics* (Oxford: Oxford University Press), 111–64.

Kirchin, Simon (2010). 'The Shapelessness Hypothesis', *Philosophers' Imprint* 10: 1–28.

—— (this volume) 'Thick Concepts and Thick Descriptions'.

Klein, Ewan (1980) 'A Semantics for Positive and Comparative Adjectives', *Linguistics and Philosophy* 4: 1–45.

Larson, Richard, and Peter Ludlow (1993) 'Interpreted Logical Forms', *Synthese* 96: 305–55.

Ludlow, Peter (1989) 'Implicit Comparison Classes', *Linguistics and Philosophy* 12: 519–33.

—— (2008) 'Cheap Contextualism', *Philosophical Issues* 18: 104–29.

McDowell, John (1981) 'Non-Cognitivism and Rule-Following', in Stephen Holtzman and Christopher Leich (eds.) *Wittgenstein: To Follow a Rule* (London: Routledge and Kegan Paul), 141–62.

Nozick, Robert (1974) *Anarchy, State, and Utopia* (New York: Basic Books).

Plato, *Republic* (several editions).

Rawls, John (1971) *A Theory of Justice* (Cambridge, MA: Harvard University Press).

Rayo, Agustín (forthcoming) 'A Plea for Semantic Localism', *Noûs*.

Recanati, François (2004) *Literal Meaning* (Cambridge: Cambridge University Press).

—— (2010) *Truth-Conditional Pragmatics* (Oxford: Clarendon Press).

Rett, Jessica (2008) 'Antonymy and Evaluativity', in Masayuki Gibson and Tova Friedman (eds.) *Proceedings of SALT (Semantics and Linguistic Theory) 17* (Ithaca, NY: CLC Publications), 210–27.

Roberts, Debbie (2011) 'Shapelessness and the Thick', *Ethics* 121: 489–520.

Scheffler, Samuel (1987) 'Morality Through Thick and Thin: a Critical Notice of *Ethics and the Limits of Philosophy*', *Philosophical Review* 96: 411–34.

Soames, Scott (2008) 'Drawing the Line Between Meaning and Implicature–and Relating Both to Assertion', *Noûs* 42: 440–65.

Sturgeon, Nicholas L. (1991) 'Contents and Causes: A Reply to Blackburn', *Philosophical Studies* 61: 19–37.

Sundell, Tim (2011) 'Disagreements about Taste', *Philosophical Studies* 155: 267–88.

van Rooij, Robert (2011) 'Measurement and Interadjective Comparisons', *Journal of Semantics* 28: 335–58.

Väyrynen, Pekka (2008) 'Slim Epistemology with a Thick Skin', *Philosophical Papers* 37: 389–412.

—— (2012) 'Thick Concepts: Where's Evaluation?' *Oxford Studies in Metaethics* 7: 235–70.

von Fintel, Kai and Anthony Gillies (2011) '"Might" Made Right' in Andy Egan and Brian Weatherson (eds.) *Epistemic Modality* (Oxford: Oxford University Press).

von Stechow, Armin (1984) 'Comparing Semantic Theories of Comparison', *Journal of Semantics* 3: 1–77.

Wasserman, Ryan (2004) 'Indeterminacy, Ignorance and the Possibility of Parity', *Philosophical Perspectives* 18: 391–403.

Williams, Bernard (1985) *Ethics and the Limits of Philosophy* (London: Fontana).