

Croatian Journal of Philosophy
Vol. XVII, No. 51, 2017

Value Disagreement and Two Aspects of Meaning

ERICH RAST

New University of Lisbon, Lisbon, Portugal

The problem of value disagreement and contextualist, relativist and metalinguistic attempts of solving it are laid out. Although the metalinguistic account seems to be on the right track, it is argued that it does not sufficiently explain why and how disagreements about the meaning of evaluative terms are based on and can be decided by appeal to existing social practices. As a remedy, it is argued that original suggestions from Putnam's "The Meaning of 'Meaning'" ought to be taken seriously. The resulting dual aspect theory of meaning can explain value disagreement in much the same way as it deals with disagreement about general terms. However, the account goes beyond Putnam's by not just defending a version of social externalism, but also defending the thesis that the truth conditional meaning of many evaluative terms is not fixed by experts either and instead constantly contested as part of a normal function of language.

Keywords: Disagreement, meaning vectors, externalism, metalinguistic negotiation, truth-conditions.

1. Introduction

Within the recent debate about relativist semantics of evaluative predicates and the corresponding notion of faultless disagreement attention has shifted towards more general discussion of value disagreement. The problem is how to account for substantive value disagreements without degrading them to merely verbal disputes. For it seems that if two people disagree about what is good in a given situation, for instance, if they associate different criteria with words like 'good' and 'better than' and one of them says 'Capitalism is good' and the other one replies 'No, it is not', then it might appear as if they only disagree about the meaning of 'good' and in the end only argue about words. In reply to this form of relativism, Alexis Burgess, David Plunkett, and

Tim Sundell (henceforth abbreviated BPS) have argued in a series of articles that such disputes are metalinguistic negotiations about the best use of a word in a context but are nevertheless substantial.¹

In this article I argue that the metalinguistic negotiation account is incomplete. BPS are right that value disagreements can be metalinguistic in the sense of being disagreements about the meaning of evaluative expressions. These meanings are not negotiated, though, nor is the ‘best use’ of a word negotiated. Rather, value disputes are instances of ordinary meaning disputes about general or abstract terms and corresponding predicates. If arguments from early semantic externalists like Putnam are taken seriously, such disputes are a normal function of language and there is no substantial difference between the way we disagree about utterances containing the word ‘electron’ and those that involve a use of words like ‘good’ and ‘capitalism’. However, Putnam’s version of externalism is not directly applicable to value disputes because of its strong externalist assumptions. Instead, a dual aspect theory of meaning is suggested that only presumes a weak form of semantic externalism for evaluative expressions and remains agnostic about the strong externalism thesis for natural kind terms. According to my suggestion, speakers often disagree about what an expression *really* means, about what I call its *noumenal meaning*, on the basis of a shared but possibly incomplete *core meaning*. Only mastery of the core meaning is required by virtue of linguistic competence.

The remainder of this article is structured as follows. In Section 2 contextualist, relativist, and metalinguistic theories of value disagreement are laid out. After a brief critique of the metalinguistic negotiation view I introduce the two aspects of meaning mentioned above and show in Section 4 how they explain metalinguistic value disputes. The result should be understood as a précis of the approach of BPS and is defended against possible objections in Section 5. A short summary is given in Section 6.

2. Contextualist and Relativist Disagreement

In this section some forms of disagreement are laid out on the basis of distinguishing what two discourse participants disagree about in a traditional truth-conditional approach: Is it the meaning of words, their truth-conditions, or various features of the context and/or the circumstances of evaluation of an utterance?

In order to address this question something must be said about truth-conditions first. As young Wittgenstein put it: “To understand a proposition means to know what is the case, if it is true.” (Tractatus, 4.024)² In modern versions of truth-conditional semantics it has

¹ See Sundell (2011); Burgess (2013); Burgess and Plunkett (2013); Plunkett and Sundell (2013, 2014).

² Cit. in Wittgenstein (1969).

become customary to additionally provide a way to deal with indexical expressions such as the tenses, and in the philosophy of language two main traditions have become prevalent. In a Lewisian account based on (Lewis 1980) a sentence φ is true or false in a model relative to a context c and an index i . The context provides a way to fix the denotation of indexicals, whereas the index consists of world-time pairs derived from the time and world of the context.³ The purpose of the index in this type of theory is purely technical: Expressions like ‘yesterday’ or ‘it is possible that’ implicitly quantify over it as part of their meaning. For example, ‘it is possible that φ ’ is true in c, i if and only if φ is true in c, i' for an index i' that is the same as i except that the world of i' is among those that are closest to those of i .⁴ Importantly, the world and time of an index is derived in some automatic way from the context.

Kaplan’s two-dimensional semantics is similar to Lewis’s suggestion in many respects.⁵ The main difference is that Kaplan’s logic is weakly intensional and two-layered, thereby introducing the notion of the semantic content of an utterance:

| | | | | |
|--------------------|---|---------|---------------|------------------|
| linguistic meaning | + | context | \Rightarrow | semantic content |
| semantic content | + | CEs | \Rightarrow | extension |

Hereby, ‘CEs’ is an abbreviation for circumstances of evaluation, the way Kaplan prefers to call the index. In order to avoid confusions with more general talk about modal indices by other authors like Cresswell (1990, 1996) this abbreviation will be used from now on. Kaplan’s approach allows for distinctions based on semantic content, which are intensions that play the same role as propositions in one-dimensional theories, and because of this increased expressivity it will be used as a basis of the following considerations unless otherwise noted. However, like in Lewis’s approach, CE’s play a merely technical role as a means to implicitly quantify over world-time pairs while keeping the interpretation of indexicals rigid. As an example, if Bob says ‘Yesterday, I was here’ in Pasadena on 5/23/2016, this is evaluated in such a way that the semantic content of ‘I is Bob, that of ‘here is Pasadena and the utterance as a whole is true if and only if Bob was at Pasadena on 5/22/2016, the day before the the day of the context.

Both theories allow for several distinct types of disagreement, some of which are *metalinguistic* and some of which are *content-based*. Starting with the latter first, consider the dialogues (1ab) and (1ac):

³ See Lewis (1980: 85–88).

⁴ In this formulation, Lewis’s neighborhood semantics is assumed. In contrast to this, in a normal modal logic the world of i' must be accessible from the world of i (= the world of c , in this case, because there are no nested modal operators) by a dyadic accessibility relation. These details make for large technical differences but we can ignore them in what follows, since the nature of the respective modalities plays no role in the following discussion.

⁵ See Kaplan (1989). He calls the linguistic meaning of an expression its ‘character’, but we stick with the term meaning.

- (1) a. Anna: Capitalism is good.
 b. Bob: No, it isn't.
 c. Bob: No, Democracy is not good.

The idea of direct content-based disagreement is that Anna expresses some semantic content φ that Bob denies directly in the sense that his reply expresses an ordinary truth-functional negation $\neg\varphi$ of the original content. He makes this explicit in (1–c). Attributing this form of disagreement only makes sense when we have reasons to believe that Anna's utterance and Bob's reply are based on the same linguistic meaning and take place within the same context of use, as the example suggests. In a Kaplanian framework the semantic content of Bob's utterance can then be taken as the negation of Anna's to the effect that he directly contradicts her.

Within these theories there are many other forms of disagreement that are not discussed in the literature very often. Two agents may disagree about empirical facts of the context of use, for example about the question who is the speaker or what the time of utterance was. Two agents may also disagree about contextual factors that are given by facts but rather belong to the cognitive context, to a pragmatic theory of interpretation, and are nevertheless often muddled into the context parameter.⁶ For example, they may disagree about the place denoted by a use of 'here' in the following utterance:

- (2) Anna: It's cold here.

Is it cold in the bed, cold in the hut, cold in Juneau, Alaska, cold in Alaska at this time of the year, cold on this continent as opposed to Australia, cold on Earth as opposed to Venus, cold in the Sagittarius arm of the Milky Way, cold in this part of the universe, and so forth? The indexical can have any of those intended meanings and the disagreement is not semantic in the sense of being disagreement about a referent that is provided by linguistically-mandated rules. Instead, the disagreement is based on different interpretations when the hearer does not take up the speaker's referential intentions. This pragmatic disagreement about the context can occur even if both speakers agree on the place of utterance. Note that this type of disagreement can also occur in the above example (1), namely about the extension of the present tense that is part of the meaning of 'is'. Anna might want to convey that capitalism is good in general (i.e., the generic reading of the present tense), but she may also intend to convey the assertion that capitalism is good for us now, as opposed to capitalism in Ancient Greece or in the near future.

Two agents may also disagree about the question whether a certain expression is context-sensitive and about which features of the context

⁶ The term cognitive context is borrowed from Penco (1999) who introduces other useful distinctions. I have laid out my own view about this type of context at several occasions, see e.g. Rast (2009, 2014).

are relevant for determining its semantic content. For example, Anna and Bob may disagree about the question whether the token of ‘here’ in (2) denotes the hut, because Anna intended the token to do so, and Bob is just wrong about taking it to stand for Alaska, or whether the semantic content of (2) is underspecified, and Bob can therefore claim that his interpretation has some legitimacy, since Anna ignored the fact that they were just talking about the climate in Alaska when she made the utterance. Whatever stance one takes about their dispute, it is hard to deny that their disagreement is metalinguistic. It is also not hard to imagine circumstances under which (1) might constitute a similar dispute *if* ‘good’ is context-sensitive as many moral philosophers would claim. For example, Anna might argue that she meant ‘Capitalism is good for me’ and this is the correct interpretation given that it was her intention to convey this message, whereas Bob might argue that ‘good’ principally has no genuine relational reading, as Moore (1903) famously claimed.

Relativist positions have been developed as alternatives to the classical contextualism outlined above, though within the same truth-conditional setting, and defended for epistemic modals, future tenses, knowledge attributions, aesthetic predicates, predicates of personal taste, and evaluative language in general.⁷ The key difference between semantic contextualism and relativism is that in the former any additional parameters of the CEs are derived from corresponding features of the context. MacFarlane (2012) calls a position that puts nontraditional factors into the context and, by some linguistically mandated derivation, into the CEs at later stages of evaluation *nonindexical contextualism*. Consider the following utterance, for example:

(3) Bob: Roller coasters are fun.

If the linguistic meaning of ‘fun’ is taken to be context-sensitive to an experiencer and this experiencer is determined in the context of use such that the semantic content ends up with a specification of this experiencer, then this is a nonindexical contextualist semantics for ‘fun’. The representation is semantic, because the additional factor is mandated by linguistic meaning and fully specified within a context. It is nonindexical, because it need not be claimed that the additional factor is determined by the context of use in the narrow sense *only*. The experiencer need not be the speaker but could rather be determined on the basis of the speaker’s intention (e.g. in the reading ‘fun for us’) or by the conversational context.⁸ The representation is contextualist, as long as the experiencer is derived or otherwise determined from the context of use. Hence, in a two-layered model the semantic content will change whenever the respective feature of the context changes.

⁷ See for instance Kölbel (2002), Lasersohn (2005), MacFarlane (2005, 2014), Egan (2014).

⁸ While there may be concerns about this way of ‘semanticizing’ possibly pragmatic interpretation processes, these need not worry us here. It is one possible position that may give rise to a particular kind of contextual disagreement.

In contrast to this, according to *assessment-relativism* contexts and CEs are decoupled from each other, and an additional assessor or relevant features of an assessor such as evaluative standards are taken to be a constitutive part of the CEs. This means in the two-layered model that one and the same utterance in a given context of use expresses the same semantic content, but that the truth or falsity of this content may vary from assessor to assessor. So if Anna disagrees with Bob about (3) by uttering ‘No, Roller coasters aren’t fun’, this may be taken as a case of faultless disagreement. (3) in the context c in which Bob is the speaker expresses some semantic content p that leaves the assessor of ‘fun’ unresolved, i.e., the proposition that roller coasters are fun. This content p then comes out true or false in a model relative to the respective features of the assessor of the CEs. In the example, if the assessor is Anna, then it will be false, and if the assessor is Bob, then it will be true. Since the assessor is not derived from the context of use, Bob’s assessment has no priority over Anna’s. Both are right, each from his or her own perspective, and a third-party observer may in turn agree with either Bob or Anna.⁹ So they disagree, as indicated by Anna’s reply, yet both of them may be right. There is leeway for some subtleties in the way assessors are assigned to CEs which may have applications in moral philosophy. Neither the speaker nor the hearer need to be the proper assessors, for instance, they could be ideal observers instead and many more refinements are possible.

3. *Pragmatic and Metalinguistic Replies*

One criticism of direct disagreement is that it is content-based and that disagreement is way more flexible than that. Two people may also be said to disagree if they have differing attitudes towards information that is conveyed pragmatically. For example, Lopez de Sá has argued that speakers may disagree in a contextualist semantics even if they associate different semantic contents with the same utterance, because at the same time they might pragmatically convey presuppositions of commonality.¹⁰ In such a theory, when it is also based on speaker-contextualism, (1–a) could be understood as expressing the proposition that democracy is good from the perspective of Anna, whereas (1–c) would express the proposition that democracy is not good from the perspective of Bob, which is compatible with Anna’s assertion, and they might still disagree at a pragmatic level, because each of them presupposes something like ‘What is good for me, is (usually) good for everyone’ and ‘What is not good for me, is (usually) not good for anyone else either.’ This peculiar type of contextualism has probably never been defended for goodness in general, but when ‘democracy’ is replaced with

⁹ If the lack of a truth value is also allowed, then a third party observer need not even agree with any of them.

¹⁰ See de Sá (2008, 2009, 2015).

'licorice' the natural reading of 'good' in this context may act like a surrogate of 'tasty' and for this reading the semantics may not seem less plausible than the corresponding relativist position according to which the content of Anna's utterance is 'Licorice is good' simpliciter, which turns out to be true when assessed by Anna and false when assessed by Bob. To cut a long story short, while there may always be a battling of intuitions about the question of whether and in which way various sorts of expressions and their readings are context-sensitive, de Sá's critique seems justified in general. It seems inadequate to presuppose a notion of disagreement that is only defined on the basis of what and how utterances express semantic content. This critique extends to the other content-based notions of disagreement mentioned above, because they are based on aspects of the semantic representations rather than attitudes of speakers about speech act content or other facts about the conversational situation or their own states of beliefs and desires that need not be directly connected to semantic representations.

A bigger problem with direct content-based analyses of value disagreement is that they link the expressions to certain metaethical positions like non-ideal appraiser-subjectivism that may or may not be appropriate for a given evaluative predicate. So even if they explain value disagreement correctly for certain expressions like predicates of personal taste, they may not be adequate for explaining value disagreement in general.

While authors like de Sá have tried to defend their own contextualist positions against relativism, Plunkett and Sundell (2013) have argued that many cases of disagreement are metalinguistic, yet can be substantive, are worth having, and are not merely verbal disputes. A metalinguistic approach is appealing for the analysis of evaluative expressions at first sight, because it allows one to explain substantive value disputes without committing to a contextualist or relativist semantics. Consider example (1) again. After at least 2,500 years of philosophical thinking about goodness, nothing can be said about the concept without taking sides, of course, but it is fair to say at least that many philosophers are willing to follow von Wright (1963) in his assessment that there are many varieties of goodness and that there is a reading of 'good' in (1) that is neither hedonic nor purely instrumental. As already suggested, under such a reading it is hard to justify a contextual or relativist semantics, for this would mean endorsing some naive form of speaker- or appraiser-subjectivism that many philosophers would consider problematic, and BPS deliver a more neutral analysis of corresponding types of value disagreement. Before turning to my own proposal, I therefore wish to briefly discuss their approach. Since their articles address many issues in semantics and metaethics at once, only the central points can be repeated here.

First, BPS lay out what they call 'canonical disputes'. Two persons A and B are in a canonical dispute if A's utterance expresses some se-

mantic ‘object’ p and B’s utterance another semantic object q such that p and q are fundamentally in conflict with each other (Ibid.: 9, paraphrased). By formulating their conditions in this general way, BPS allow canonical disputes on the basis of incompatible plans and desires, so their notion of disagreement is already broader (and, consequently, less revealing) than the ones laid out above. However, they still have similar worries as de Sá about relying solely on a semantic content-based definition of disagreement and therefore formulate the following criterion:

Disagreement Requires Conflict in Content (DRCC): If two subjects A and B disagree with each other, then there are some objects p and q (propositions, plans, etc.) such that A accepts p and B accepts q , and p is such that the demands placed on a subject in virtue of accepting it are rationally incompatible with the demands placed on a subject in virtue of accepting q . (BPS 2013: 11)¹¹

They then argue that there are examples that intuitively count as disagreements and fulfill DRCC, even though they are not canonical in the above sense, and conclude that ‘...theorists take a wrong turn as soon as they conflate the question of whether a disagreement is *genuine* with the question of how the information on which a disagreement centers happens to be communicated.’ (ibid: 13) This paves the way for their thesis that many substantial disagreements about utterances involving evaluative expressions are metalinguistic in nature and that they are nevertheless worth having. Here are their main examples:¹²

- (4) a. That chilly is spicy!
b. No, it’s not spicy at all!
- (5) a. Secretariat is an athlete.
b. Secretariat is not an athlete.¹³
- (6) a. Waterboarding is torture.
b. Waterboarding is not torture.
- (7) a. Lying with the aim of promoting human happiness is sometimes morally right. In fact it often is!
b. No, you are wrong. It is never morally right to lie in order to promote human happiness.
- (8) a. Tomato is a fruit.
b. No, tomato is not a fruit.

According to BPS, all of the above dialogues are examples of *metalinguistic negotiation*. They concern the question of how to best use words

¹¹ This requirement fixes deficiencies of earlier definitions of disagreement based on attitudes that one person cannot hold jointly like in Egan (2010: 278). Cf. Marques (2014) for a critique of such definitions. I do not wish to enter this methodological debate here and consider DRCC a reasonable rule of thumb.

¹² See (Plunkett and Sundell, 2013: 15, 16, 19, 20, 22).

¹³ Two sports reporters are discussing a horse in a race. One is calling it an athlete, whereas the others point is that only humans can be athletes. This example is originally from Ludlow (2008).

in a given context and nevertheless may be important and substantive. BPS call the business of discussing and determining how we *should* use words or concepts ‘conceptual ethics’.¹⁴ For example, in a biology class (8–a) might be the appropriate position, whereas (8–b) may be more adequate for a chef. Likewise, there is no doubt that (6) is morally relevant and a dispute worth having, even when underlying the dispute is, according to BPS, the fact that both discourse participants disagree about the meaning or adequate definition of ‘torture’. What constitutes the concept under discussion is important because of ‘...sociological facts about its sociological role’ (BPS 2013: 25), because there is something ‘...substantive at stake in how the relevant terms are used in the context [...] and the speakers recognize this fact.’ (ibid.) As they lay out, these disputes also survive paraphrasing, a test devised by Chalmers (2011) to distinguish substantive from merely verbal disputes, so metalinguistic disagreement need not be merely verbal. That is, in a nutshell, their position.

To summarize, the problem is how to explain value disagreement in a way that does not make it a mere fighting about words and without stipulating a particular theory of value. Notwithstanding the possibility that many evaluative predicates are context-sensitive in some other ways—in fact, most of them are—, explaining examples like (1) and (6) in the contextualist way does not seem to make them cases of genuine disagreement. de Sá (2008, 2015) argues that this apparent deficiency can be fixed by explaining the disagreement by different attitudes about pragmatically conveyed content. However, relativists are not satisfied with this solution, as the contextualist approach looks like an attempt to explain away the strong intuition that the discourse participants in such examples really do talk about the same subject matter. But the relativist position presupposes a type of appraiser-subjectivism that is even stronger than contextualism, and both positions seem implausible as a general way of explaining examples like (1)–(8). They may be adequate for *some* expressions such as ‘tasty’ and ‘fun’ but seem to be unacceptable for other, broader value predicates like ‘brilliant’, ‘right’, ‘is torture’ and readings of ‘good’ that are not purely hedonist or instrumental.¹⁵ The problem is that sometimes people might even disagree fundamentally about the underlying value theory, for example in (7) the speaker might be a consequentialist and the hearer a Kantian. Following de Sá in the assessment that the discussion between contextualists and relativists was based on a too narrow semantic view

¹⁴ See (ibid: 3) and Burgess (2013); Burgess and Plunkett (2013).

¹⁵ To avoid misunderstandings, it must be stressed that relativists generally do not deny this and also disagree among themselves about the question which expressions have a relativist semantics. In the same vein, it is also not claimed in the next section that a dual aspect semantics invariably means that no expression has a relativist semantics. Some might very well have a meaning that is best analyzed in relativist terms.

of disagreement, BPS suggest that many cases of value disagreement are potentially important and worthwhile metalinguistic negotiations.

4. *Two Faces of Meaning*

This section consists of a negative and a positive part. In the negative part, I argue that the metalinguistic approach is not satisfying as an explanation of the value disputes under consideration. In the positive part, I sketch a dual aspect semantics that fixes these shortcomings and can explain disagreement in a way that is content-based but nevertheless remains compatible with the thesis that many value disputes are about the meaning of the expressions involved. This suggestion is then defended and motivated in more detail in Section 5.

4.1. *Why Metalinguistic Negotiation Does Not Suffice*

Overall, the metalinguistic approach has many merits, as it paints a more realistic picture of some value disputes than purely semantic notions of disagreement. However, the problem is that it leaves the question open of how to define metalinguistic negotiation and it turns out at a closer look that it is very hard to make sense of this process. According to BPS, metalinguistic disagreement hinges on the idea that ‘...certain words (largely independent of which specific concept they express) fill specific and important functional roles in our practices[.]’ (Plunkett and Sundell, 2013: 20) This passage seems to suggest that the functional role of the expression in question does not depend substantially on the concept it expresses, but is this really plausible? To me it is not, for it seems hard to find a way in which a social practice with regards to a term may come into being without being based on a widely accepted meaning of that term, or in other words, *because* the term has that specific meaning and not another one. At other places, BPS explain that two discourse participants who disagree ‘...each advocate a view about which concept is best suited to play a certain functional role in thought and practice...’ (ibid.: 21), and so it seems that the disagreement is about the concepts again. Perhaps the question whether the disagreement is about the use of the words or the concepts is not so important, because according to BPS both come as a package. Anna and Bob both try to advocate their concept in association with a given word, arguing that the respective concept better suits or fits the functional social role of past uses of the word. The problem with this suggestion is, however, that BPS do not provide a satisfying account of the ‘...social, historical, and psychological facts about what is standardly associated with the use of that term.’ (ibid.) To see why this is a problem, consider the following joking conversation:

- (9) a. Anna: Work is torture.
 b. Bob: Indeed, it is.

They are joking, but why? The correct answer is in my opinion that they both know that in modern democratic societies without slave labor camps work does not classify as torture, therefore Anna's utterance is obviously false, and then a Gricean recovery strategy kicks in.¹⁶ This can only work if 'torture' has enough of a partial pre-established meaning on the basis upon which Anna and Bob can agree that work does definitely not qualify as torture under normal circumstances, even if they disagree about other torture-related issues such as (6). For the dialogue to work as a joke, a pre-established meaning is responsible and *not* the fact that 'is torture' is not commonly used in a way that an utterance of 'Work is torture' is accepted. Otherwise a value disagreement like (1) or (6) would boil down to an argument not about what constitutes torture, but about how people in the past used the word—but past utterances of (6–b) might not have existed at some point in time or they might have been made by people who clearly agreed all the time that (6–b) is false. So it seems that there is a pre-established meaning on the basis of which Anna and Bob forward their own views about torture, where one or both of them might deviate from this meaning, which past uses have expressed. If so, a dispute about past uses is in the essence a dispute about this pre-established meaning. However, it seems woefully inadequate to reconstruct Anna's and Bob's dispute as being about the question whose definition of torture best fits this pre-established meaning. *That* question should not be hard to answer and there ought not be much rational disagreement, because it is a purely factual matter. Either her suggestion fits or does not fit the prior established meaning, and the disagreement would in the end be about a linguistic matter, whether other speakers in the community use the word 'torture' in a way that is compatible with, suggests, or even implies that waterboarding is torture. There is nothing to negotiate then, they could just ask other speakers, so this cannot be the right approach either.

Perhaps it helps to understand the social functional role in a very broad sense and to assume that Anna and Bob disagree about a case that is not settled by prior uses. In this view, they seek to *extend* the meaning of 'torture' in a way that harmonizes with prior social practices concerning torture such as banning it, persecuting it as a crime, and so forth. However, this cannot be right either. Even if everybody had always agreed that there is no torture in some speaker community, it is hard to see how this fact alone could settle the question or even just play a substantial role in determining whether waterboarding is torture. Suppose Anna and Bob continue to discuss (6) and finally come to agree that waterboarding is torture. According to theory that BPS seem to suggest, Anna would essentially have convinced Bob of the fact that claiming that waterboarding is torture best suits our previous

¹⁶ This is only an example. There are other explanations such as those given by Relevance Theory and nothing in what follows hinges on assuming the classical Gricean picture.

practices of using ‘torture’ and reasonably fits with the already existing sociological functional role of the word. But if Anna and Bob negotiate in that way, they still do not settle the question whether waterboarding *is* torture! They could both agree that waterboarding is not torture and be wrong about it. How so, if they are just negotiating? It seems that the answer is this: They really argue about what constitutes torture in (6), about what makes socio-economic systems good or bad in (1), and about what’s morally right in (7), and in each case they argue on the basis of some already shared, pre-established meaning but without necessarily presuming that this meaning is correct or adequate.

This critique does not concern the metalinguistic aspect of the proposal. Dialogues like (1)–(8) are metalinguistic in the sense that they concern the meaning of expressions involved, and it seems clear enough that these expressions are not used in the ordinary way in these dialogues. I fully agree with BPS about this observation. The critique challenges the view, however, that Anna and Bob negotiate the best or most appropriate use of a term. The thesis that in (6) Anna suggests to Bob that waterboarding should be called ‘torture’ because of existing practices associated with the word is not very convincing as a *general* explanation of such disputes, because this would make her use of ‘torture’ a mere proxy or surrogate for something else. Instead, the correct inference goes the opposite direction: Since the word expresses a certain concept, it is endowed with a certain social function, and another use of the word is endowed with the same (or a more specific) social function *insofar as* whatever it expresses falls under this concept. Hopefully, in Anna’s society torture is morally prohibited, illegal, and socially unacceptable, and she may also use (6–a) to implicate or otherwise pragmatically convey these facts with the intention of laying out to Bob that waterboarding should be illegal, socially unacceptable and prohibited. But even if she does so, a use of (6–a) to successfully convey this additional speech act content can only be based on the prior meanings of the words in the utterance.¹⁷ If waterboarding is not torture in the first place, as Bob believes, then how could (6) serve to indicate or even justify that waterboarding should be illegal? Alice’s reasoning chain must thus go in the following direction: Waterboarding is illegal, socially unacceptable, and so forth, because it really *is* torture, and what is considered torture in her society is considered illegal, unacceptable, and so forth. If that is the right direction of explanation, then the conundrum remains: If, on the one hand, that meaning is the pre-established meaning, then the dispute is factual, as laid out above. If, on the other hand, it is Anna’s meaning as opposed to Bob’s favorite definition, then the disagreement is once more direct and content-based, hence the two of them are talking past each other and the relativist critique applies.

¹⁷ Radical pragmatists and meaning eliminativists would not subscribe to this view. These positions discard with truth-conditional semantics altogether and cannot be taken into consideration here for lack of space.

Similar points can be made about the other examples. However, these considerations remain compatible with the view that metalinguistic negotiation sometimes takes place in dialogues like (1)–(8), they merely indicate that BPS’s approach is not adequate as a *general* explanation of these kinds of examples. Doubts remain about their use of the term ‘negotiation’, which seems to be too weak for an explanation of at least some cases of important value disagreement, namely those in which the social function of an expression arises or is even constituted from its meaning. The role of negotiation in the metalinguistic approach needs to be motivated further. As it stands, there is a gap between the pre-established meaning of an expression and its social function. If the latter is determined or significantly constituted by the former, then one of the disagreeing discourse participants may either be right by virtue of this meaning, if the new use is ‘covered’ by it, and the other will be wrong (or, both are wrong). This is a factual matter. The second possibility is that the new use is not covered by the prior meaning, as could be argued for the ‘athlete’ example (5), for instance, but then the extension of the new meaning is not covered by the social function either and the existing social function of the expression alone does not generally provide a reason for extending the prior meaning to the new use. BPS are right in calling the process of arguing for such an extension negotiation, as this process is oddly idle: The evaluative predicate has no power of its own, its existing truth-conditions barely matter, and it merely seems to act as a surrogate for attributing properties of the social function to the logical subject of an assertion. This is implausible. Although there may be exceptions to this rule, generally speaking linguistic expressions have a social role because of their meaning, not their meaning because of an existing social role.

None of the above implies that metalinguistic disagreement cannot sometimes occur, and BPS have argued convincingly that at least some such disagreements can be substantial and worth having. The following dual aspect approach should therefore be taken as an addition to their account, as a way of making it more precise.

4.2. *Core Meaning and Truth Conditions*

To address the above mentioned shortcomings semantic theory itself needs to be modified. This makes the proposal more controversial, of course, but it also becomes more expressive and the modified theory also explains many more phenomena than value disagreement. In the proposed dual aspect theory (DAT), every expression of a language has two meanings. The *noumenal meaning* of an expression contributes to the truth-conditions of a sentence as a whole in accordance with Wittgenstein’s dictum. In contrast to this, the *core meaning* of an expression represents what I have so far called pre-established meaning.¹⁸

¹⁸ There is an older use of ‘dual aspect theory’ in the philosophy of language, which is based on the idea of enriching meaning representations of sentences

This meaning is best understood by a loose bundle view, it represents the smallest common denominator between speakers of a given sociolect insofar as the meaning of the given expression is concerned, and speakers of a linguistic community have to agree to a large extent on this meaning for communication to succeed and in order to count as a competent speaker. Both meanings may be truth-conditional, because the attribute ‘truth-conditional’ marks a *method of specifying certain aspects of meaning*. Moreover, both types of meaning may be spelled out in contextualist terms in the way laid out above to deal with indexicals and other forms of overt context-sensitivity, and this will be assumed from now on without further mention.

There is a third element that for lack of a better term shall be called *noumenon*. It is the (stipulated) object of an intentional stance of speakers towards reality; noumenal meanings aim at the noumenon, they may pick out objects in reality like Fregean senses aim at the ‘Bedeutung’, and sometimes reality also *informs* speakers to revise the noumenal meaning of an expression. Noumenal meaning is also a précis of the core meaning, whereas the core meaning is first and foremost what is needed in everyday conversations. Core meaning evolves primarily out of a need to communicate and coordinate behavior, it may be conventional in the sense of Lewis (1969) or may be more broadly conceived conventional in the sense of being based on non-inferential behavioral regularities between word and meaning. Noumenal meaning generally evolves out of core meaning when people start asking questions that aim at reality. For example, two speakers can talk about a lightning during a thunderstorm and agree that it needs to be avoided and that it is likely that fire can be found where the lightning has struck. But they may also start asking themselves what a lightning really is. Is it a sign from the gods, some special form of fire, an electromagnetic phenomenon? These questions concern the noumenal meaning.

Core meaning is often incomplete and loose. For example, ‘freedom’ is an abstract noun that in everyday use stands for the possibility of making voluntary choices, an absence of oppression and unnecessary prohibitions, a lack of restrictions in thinking and acting, et cetera, but different speakers will only loosely agree on such features. In contrast to this, the scientific definition of an expression by an expert in the respective field is a possible candidate for a noumenal meaning or an approximation thereof. Kant’s definition of ‘freedom’ is such a candi-

occurring in attitude ascriptions with additional features (guises, cognitive roles, ways of being given, Fregean senses, etc.) in order to explain cases of referential opacity. For example, Lois Lane might believe that Superman can fly while at the same time consistently believing that Clark Kent cannot fly, because Superman is given to her in cognition in a different way than Clark Kent. Core meaning does not fulfill this role, it serves the opposite goal of standing for the lowest common denominator and is therefore looser than noumenal meaning rather than more fine-grained. The two theories have nothing particular in common. To combine them, a triple aspect theory would be needed.

date, for example. Importantly, both meanings may be revised on the basis of a change to the other. The third element, the noumenon, is not a sort of meaning but something else in reality such as a mathematical fact, the orbit of the Earth around the sun, a particular family of cats, a building, or a person, if the noumenon exists at all. Since it is not a sort of meaning, the account is a dual rather than a triple aspect theory.

Before defending DAT in more detail, let us take a look at how it fares with the problematic examples. When two speakers disagree like in (1)–(8), they (mostly) agree about the core meaning and at the same time disagree about the noumenal meaning. They also *intend* the noumenal meaning to match some aspect of reality.¹⁹ So what they really disagree about is the question whether *their* respective candidate for a noumenal meaning is adequate with respect to the stipulated noumenon and can therefore be regarded as a précis of the imprecise everyday core meaning on which they both implicitly agree. If one of the discourse participants does not associate a core meaning with an expression that is sufficiently similar to that of other members of the linguistic community, as they evidence from that speaker's odd linguistic behavior, then they will raise doubts about his competence as a speaker of the respective sociolect. Core meaning is linguistically mandated and the basis of a shared lexicon. For example, it is part of the core meaning of 'torture' that it involves inflicting serious harm (often though not always in an attempt to extract information), it is part of the core meaning of 'athlete' that an athlete is physically fit under normal circumstances, and so on. If Bob claimed that torture does not harm the person who is tortured, then he would not be considered a competent speaker with respect to 'torture' and Anna would react accordingly, perhaps by shaking her head in disbelief or by trying to educate him about the English language. In contrast to this, competent speakers may disagree about the noumenal meaning at any time, it is principally contestable and often hinges on complicated background theories which might turn out to be false. Disagreement about noumenal meaning is never about this meaning alone, though, which would lead to similar problems as those laid out for metalinguistic negotiation in the previous section, but rather about the question whether this meaning adequately captures the respective aspect of reality it is intended to capture—whether it correctly points to the noumenon.

From this perspective, the disagreement between Anna and Bob is metalinguistic insofar as it concerns an aspect of the meaning of expressions used, but it is also linguistic and a common form of disagreement, because it is based on a mechanism that is completely normal for language users and occurs daily. Linguistic competence does not require convergence on noumenal meaning, on the contrary we often discuss it when we talk about general terms that are not grounded in

¹⁹ Whenever intentions are mentioned in this article, they are to be understood in the philosophical sense that started with Brentano, not as intentions to act.

empirical facts in any obvious way. In such discussions, expressions are not mentioned in the sense of being quoted, but the underlying concepts are challenged by one or more discourse participants when they advocate some particular noumenal truth-conditions for the term. This may, but need not be an argument for or against a relativist or contextualist semantics, as (3) and (4) might exemplify, because every aspect of the noumenal meaning is open for discussion. Some speaker might argue for the assessor-sensitivity of a given evaluative predicate, for example. A discussion may also amount to a mostly verbal dispute, as one might consider Ludlow's 'athlete' example (5), for it is questionable whether there is something in reality that can decide whether horses can be athletes or not other than the existing core meaning or arbitrary conventions.²⁰ In yet other cases such as (8) the dispute might be best characterized as a linguistic misunderstanding of confusing contextual standards, and this is also how BPS discuss this example. In a biology class the noumenal meaning is decisive, since biologists study plants and know criteria for classifying tomatoes that ordinary speakers might not even have heard of. If Bob is not aware of this context and defends the core meaning, he is right, too, but in a sense that is often accompanied with a sigh. He is wrong in the same sense as a student in a philosophy class would be ill-advised to use 'realist' in the sense of 'an uncomplicated and reasonable person, being down to earth'.

Some such cases might be conflicts between different sociolects, each equipped with its own noumenal and core meaning for a term, while others may concern the context of use, whether the word is used in an every day sense, or whether we are interested in *what it really means*. But this inaccuracy is not a deficiency of DAT, it merely reflects the reality of disagreements that everybody has experienced before. What is important for my thesis is that in clear cases of indirect value disagreement like (1), (6) and (7) the dispute will normally be about the noumenal meaning *on the basis of the idea that something in reality corresponds to this meaning and thereby justifies it*. Provided that both suggestions are sufficiently compatible with the core meaning established by the linguistic community, such a dispute will be merely verbal and insubstantial only if (a) competent speakers do not intend the noumenal meaning to 'fit' the noumenon *and* (b) there is nothing in reality that would justify the meaning. Notice that error theorists defend (b) but need not deny (a). As Meinong (1971) argued pervasively, having an intention in the sense of the Brentano School need not imply that the intentional object exists.²¹

²⁰ Bear in mind that many disputes may also be about the core meaning; speakers are not always a hundred percent competent, the boundaries between sociolects can be vague, and it may be disputed which sociolect is relevant to the discussion.

²¹ I have laid out my own stance about nonexistent objects in Rast (2011). However, abstract objects are more complicated, see Zalta (1983) for a logically and philosophically developed approach.

Regarding value disagreement, the theory explains how speakers may disagree about the meaning of expressions without talking past each other in the sense of direct contextualist disagreement and without assuming a relativist stance that may only be adequate for select predicates of personal taste. Principle DRCC is fulfilled, because the speakers disagree about the noumenal meaning of an expression on the basis of a given core meaning *and* and on the basis of their beliefs about the noumenon. The theory is therefore an extension of the metalinguistic approach, which on its own has problems with explaining why and what people are negotiating about the meaning of an expression. However, if this was the only merit of DAT, then the metalinguistic approach would perhaps be more desirable, since any unnecessary duplication of entities should be avoided by application of Occam's Razor. As I will argue in the next section, DAT has many additional benefits and is nearly unavoidable for semantic externalists anyway.

5. *In Defense of DAT*

The dual aspect approach laid out in the previous section is not new. Similar suggestions have been made by semantic externalists and 'anti-individualists' like Putnam (1975b) and Burge (1979, 1986) on the basis of insights from Kripke (1972). However, some care is needed to keep different aspects of externalism apart, since otherwise we would only win a pyrrhic victory. DAT aims at remaining compatible with weak semantic internalism by focusing on the social aspect of linguistic labor division and allowing at the same time that two disagreeing parties may have different noumenal meanings *in mind*. In contrast to this, the strong externalism of Putnam (1975b) stipulates that a description of the extension constitutes part of the meaning of a term and that knowledge of this description or of the extension itself are not required by virtue of linguistic competence. The problem with this view is that it seems fairly implausible as a requirement for moral terms, because it presupposes an implausible form of moral realism, and this view would also be incompatible with more recent work in Putnam (2002). I will argue for a weaker form of externalism that explains metalinguistic disputes by the fact that speakers disagree about the noumenal meaning of expressions and in order to satisfy DRCC speakers must somehow be aware of these meanings in their minds in case of a genuine disagreement. This weak externalism for value terms remains compatible with the strong externalism of natural kind terms and nothing in what follows implies that strong externalism needs to be given up for those terms.

5.1. *Arguments from Social Externalism*

All of the arguments for semantic externalism cannot be repeated here for lack of space. However, I want to sketch two of them very briefly in

order to carve out the distinction between weak and strong externalism I am aiming at. There are two types of arguments for semantic externalism. The first one is based on Twin Earth scenarios, whereas the second one pertains to social labor division and is empirical. According to the empirical thesis, individual speakers of a natural language like English may generally be judged as competent speakers, even though they do not, according to the judgment of experts, ‘associate’ or ‘grasp’ the truth-conditional contribution of an expression to a sentence-in-use as a whole precisely enough.²² For example, speakers of English can use ‘birch’ competently without being able to distinguish birches from elms, can use ‘water’ without knowing that it is H₂O, and can use ‘capitalism’ competently without implicitly knowing or being able to successfully apply a definition that would satisfy an economist. The empirical evidence for social externalism is pervasive and there can hardly be any doubt that there is a linguistic labor division. We leave the question of what exactly certain expressions mean to experts and generally assume that these experts have a precise idea about what these expressions mean. Empirical arguments for social externalism are compatible with semantic internalism, though they certainly make it less appealing. An internalist could claim that at least some expert has to implicitly know the ‘right’ meaning of an expression in order for it to make sense at all; if that were not so, the internalist could argue, sentences containing such expressions might not be about any common subject matter at all.

Putnam’s Twin Earth argument closes this loophole. Suppose John on Earth has Twin John on Twin Earth as a counterpart who is molecule-to-molecule identical to John and also identical in his mental life with him, except that he lives on Twin Earth where XYZ replaces H₂O.²³ Putnam argues that even though they are in identical mental and physical states, in this scenario ‘water’ uttered by John on Earth refers to H₂O, whereas ‘water’ uttered by Twin John on Twin Earth refers to XYZ. Hence, the meaning of ‘water’ is not in their heads. Similar arguments can be made for other natural kinds like ‘tiger’ and ‘gold’. These arguments are clearly not empirical, since otherwise the existence of Twin Earth and XYZ would have to be established for the argument to be sound. Instead, they are thought to reveal the principally indexical nature of certain expressions: The extension of ‘water’ is fixed indexically and possibly independent of our current state of knowledge. If the word is used on Earth, then it refers to H₂O. If it is used on Twin Earth, then it refers to XYZ.

²² The scare quotes around ‘associate’ and ‘grasp’ are certainly adequate in this context. Ideally, much more would be said about these problematic metaphors but for lack of space this issue has to be suppressed. I will continue using the verb ‘associate’, which may be taken to stand for a form of implicit knowledge or some ability.

²³ It is sometimes argued that John and Twin John could not be identical, because the latter partly consists of XYZ instead of H₂O. While technically correct, this argument leads astray and ultimately does not invalidate the thought experiment.

Since it would be implausible to stipulate that speakers of a language must know the correct extension of each term and this would also conflict the empirical observation of social externalism, and since the idea that a systematic error-theory according to which speakers are generally not competent is equally unappealing, Putnam suggests to represent the meaning of general terms like ‘water’ or ‘tiger’ by a vector that contains semantic and syntactic markers such as ‘count noun’ and ‘natural kind term’, the specification of a stereotype that comes close to what I have called core meaning, and a description of the term’s extension, i.e., H₂O for ‘water’ on Earth, but XYZ for ‘water’ on Twin Earth.²⁴ A competent speaker must only sufficiently master the first two items in a meaning vector, whereas the description of the extension is relevant for the externalist determination of meaning that Twin Earth scenarios are supposed to establish.

Putnam’s position is very close to the proposed dual aspect theory. The core meaning may be taken as being comprised of the stereotype and semantic and syntactic markers, whereas what Putnam calls a ‘description of the extension’ corresponds to noumenal meaning. However, it differs from the current proposal with respect to the role of an extension (if there is one). Putnam (1994) meanders between different readings of the third elements in his meaning vectors; he either takes it to be a *description* of the extension or as the extension itself.²⁵ In contrast to this, I would like noumenal meaning to be more broadly understood as a refined meaning with which informed speakers come up when they are intending to narrow down and state more precisely a purported aspect of reality that is *suggested* by a preestablished core meaning and sometimes, though not always, also *perceived* or *measured* from observing nature. An important difference between these views is that in DAT the core meaning of predicates must enable the language user to identify entities falling under them—or, in case of non-empirical general terms such as ‘triangle’ and ‘democracy’, to identify instances of the reified abstract objects they express—in a sufficiently precise way whenever the noumenal meaning is under discussion. This process of identification need not be foolproof, but it needs to be precise enough to ensure that two people roughly talk about the same issue. Take ‘democracy’, for instance. While the concept of democracy is characterized to some extent by examples, I submit that it cannot be defined in a purely ostensive manner. But in order for two discourse participants to discuss democracy, according to DAT they need to roughly agree on relevant aspects of the core meaning prior to being able to successfully discuss the noumenal meaning, and this involves being able to tell a democratic from an obviously totalitarian society. This prior identifica-

²⁴ See Putnam (1994: 252–255; 270–2), cf. Putnam (1975a: 148–152). Accounts of the nearness of semantic representations to each other and of individuals falling under them from a prototype have been made precise in geometrical meaning theories by Rosch (1975, 1983), Gärdenfors (2000) and others.

²⁵ See Putnam (1994: 270–1).

tion step is missing in Putnam's stereotype theory and even explicitly denied as a requirement for empirical terms, but it is crucial for understanding value disputes. According to DAT two speakers are talking past each other and cannot disagree about the noumenal meaning in any fruitful way as long as they do not have at least a glimpse about the noumenon already.

5.2. *The Peculiarities of Value Disagreement*

Externalism only *indirectly* supports the explanation of value disagreement that was tentatively outlined in the previous section. As BPS point out convincingly, it would be implausible to claim that experts generally fix the meaning of value terms, as there is persistent disagreement among experts about the meaning of particular uses of 'good' and other evaluative predicates.²⁶ Some forms of goodness are almost purely factual, however, and for these social externalism seems plausible. To this category belong medical goodness (healthier than, better with respect to health) and varieties of instrumental goodness such as what constitutes a good hunting knife or a good telescope. For a given purpose, such questions can be answered by experts, and these notions of goodness are based on criteria upon which experts concur at least to some extent. It is doubtful, however, whether the question of what constitutes a good life can be answered authoritatively by an expert, there is no agreement between those experts, and there is even disagreement about who counts as an expert in this domain. Similar doubts are in place about moral goodness, since otherwise skeptical arguments like that of Mackie (1977) could not have gained traction. So social externalism is only plausible for some limited forms of instrumental goodness, and the claim that these constitute other forms of goodness can only be made from an already assumed perspective, for example from the perspective of consequentialist utilitarianism, and not in general.

An even bigger problem occurs when strong externalism of value terms is presumed. This position is inherently problematic for evaluative language because it seems to presuppose value realism. Perhaps we point out instances of good things and actions, somehow access real values by these practices, and thereby define what a particular reading of 'better than' means, but regardless of how plausible one finds this theory, it remains just one of many possible explanations of the meaning of value terms. Lexical semantics cannot decide which moral theory is correct. Strong externalism works for empirical terms, because the assumption of metaphysical realism, the assumption that we all perceive the same actual world that is such-and-such, is philosophically acceptable to many if not most contemporary thinkers. Assuming a similarly strong realism for value terms is way more problematic.

²⁶ See Plunkett and Sundell (2013: 26–8).

Firstly, it would exclude a variety of metaethical positions, and secondly, strong arguments have been devised against forms of value realism that are based on strong externalism.²⁷

So there are tensions between arguments for semantic externalism and plausible views about the meaning of value terms. The support is indirect, since both types of arguments for externalism count in favor of DAT, which in turn explains the problematic cases of value disagreement discussed in the previous section. In order to make this story convincing, though, more has to be said about the way in which DAT explains value disagreement. To see how DAT fares in comparison, consider the moral dispute in (6) again.

Why do the discourse participants continue to disagree at all, why can't they just stipulate that waterboarding is torture₁ but is not torture₂ and happily agree with each other thereafter? According to the story of BPS, this is so, because of existing practices associated with 'torture' that are morally relevant such as condemning it, making it illegal, and so forth. However, this story is incomplete, because the discourse participants could continue to discuss whether these practices should be associated with 'torture₁' and 'torture₂' respectively, and nothing of value, it seems, would have been lost by making this terminological clarification. Why would they need to negotiate one single use and meaning of 'torture'? DAT answers this question and thus provides the missing link. The discourse participants are arguing about the noumenal meaning of 'torture', what *really* constitutes torture, on the basis of a prior partial agreement about a core meaning, which involves criteria like causing substantial harm to the victims. They disagree about the right way of making these potentially incomplete and loose criteria precise. Should psychological harm be included? How much of it? What constitutes substantial harm? What do both moral and legal experts say about it? What do international humanitarian conventions and human rights say? These questions have been discussed for some time.²⁸ From the perspective of DAT, the speakers do not introduce mere definitions for different sets of criteria, because the noumenal meaning of

²⁷ Horgan and Timmons (1992) devise a Moral Twin Earth scenario. If on Earth 'good' has a consequentialist meaning and on Twin Earth it has a deontological meaning, then we would not say that 'good' on Twin Earth and 'good' on Earth have different meanings like in the case of water, but according to Horgan & Timmons we would say that earthlings and twin earthlings have different theories of goodness (and likewise for 'right'). One need not share their intuitions and may also have doubts about the way they formulate semantic positions like Putnam's, but in any case a strong externalist of moral terms has to address this argument. In contrast to this, from the perspective of DAT earthlings and twin earthlings simply disagree about noumenal meanings and there is nothing special about the scenario that would set it apart from other disagreements about value terms.

²⁸ It may be argued convincingly, though, that the 'pro-waterboarding' side has always considered the practice torture and just flat-out lies for the purpose of strategic maneuvering. So perhaps this is not the best example, but I'd like to stick with it because BPS use it.

‘torture’ is intended to capture some aspect of reality just like the word ‘electron’ does—social, human, legal and moral reality in the former case, physical reality in the latter.

The use of ‘good’ in (1) is slightly different than the waterboarding example and indicates a feature that might be peculiar to value disagreement. Suppose Anna has an understanding of ‘good’ in mind that could be paraphrased as ‘good for me and my business’, since she is a business owner and does not aim at comparing different conceptions of society. Suppose that Bob understands it differently as ‘good for society as a whole’ and happens to be a die-hard communist. The criteria they associate with the words could not be more different, yet Anna might actually *agree* with Bob about his assessment. She just has another reading of ‘good’ in mind. The difference to the previous example is in this case that ‘good’ and its more fundamental comparative counterpart ‘better than’ are *also* context-dependent and have many different noumenal meanings as acceptable précis of the same underlying core meaning in different contexts. So in addition to disagreement about a particular noumenal meaning, competent speakers may also disagree about the way in which a use of these predicates ought to be understood within a given conversational context even if they agree on the available readings. This might be a peculiarity of what I call *aggregative value expressions* to which ‘good’ belongs. These are multi-dimensional expressions that combine various evaluative aspects of two or more items under discussion into an overall evaluation. Although I do not wish to make the case for this position here, I do believe that all genuine value expressions are aggregative in this sense.²⁹ In contrast to this, we could call ‘torture’ value-laden, because the vast majority of mentally sane persons consider it to be bad, but at the same time might not consider it a value term itself. Perhaps this is the reason why we say that torture is bad but do not often explicitly state that being healthy is good or that being brave is good: The adjectives ‘healthy’ and ‘brave’ are themselves value terms that take part of an analysis of respective medical and instrumental varieties of goodness. If this is the right view, then we may expect candidates for the noumenal meaning of value terms to be fairly complex and subject to particularly long-lasting controversies. Pace Mackie (1977), widespread disagreement about value terms need not be explained as an inference to the best explanation that there is no noumenon. Maybe a better explanation is that the disagreement is so persistent because value terms concern complex social matters whose evaluation requires considering a vast number of factors with potential trade-offs between them.

5.3. *Core Meaning and Prescriptivism*

In his famous book on evaluative language, Hare (1952) understood his universal prescriptivism as a corrected and expanded version of the much criticized Paradox of Analysis by Moore (1903):

²⁹ Stojanovic (2015) sketches such a semantic approach. Pragmatic theories are briefly addressed further below.

Moore thought that he could prove that there were no such defining characteristics for the word 'good' as used in morals. His argument has been assailed since he compounded it; and it is certainly true that the formulation of it was at fault. But it seems to me that Moore's argument was not merely plausible; it rests, albeit insecurely, upon a secure foundation; [...] Let us, therefore, try to restate Moore's argument in a way which makes it clear why 'naturalism' is untenable[.] (Hare 1999: 83–4; Sec. II.5.4)

Among Hare's arguments, the cannibals/missionary thought experiment deserves special mention, not only because it is discussed by BPS and others, but for our purpose also in order to show why Hare's preskriptivism and similar positions are fully compatible with DAT. He writes:

Let us suppose that a missionary, armed with a grammar book, lands on cannibal island. The vocabulary of his grammar book gives him the equivalent, in the cannibals' language, of the English word 'good'. Let us suppose that, by queer coincidence, the word is 'good'. And let us suppose, also, that it really is the equivalent—that is, as the *Oxford English Dictionary* puts it, 'the most general adjective of commendation' in their language. If the missionary has mastered his vocabulary, he can, *so long as he uses the word evaluatively and not descriptively*, communicate with them about morals quite happily. They know that when he uses the word he is commending the person or object that he applies it to. The only thing they find odd is that he applies it to such unexpected people, people who are meek and gentle and do not collect large quantities of scalps: whereas they themselves are accustomed to commend people who are bold and burly and collect more scalps than the average. (Hare 1999: 148; Sec. II.9.4; orig. emph.)

This much-cited passage occurs late in the book after Hare has already made the case for the commendatory nature of value terms and suggested to analyze them as sort of hidden imperatives. Its purpose is to illustrate why the descriptive meaning of value terms, whose existence Hare does not deny, are secondary to their evaluative use as tools of commendation. The missionary and the cannibals can understand each other, because they associate the same evaluative meaning with 'good'—that of general commendation—, even though they associate opposing descriptive criteria for what constitutes a 'good person' with the word.

From the perspective of DAT, such examples can be explained by stipulating that part of the core meaning of positive value terms like 'good' is commendation, and part of the core meaning of negative value terms like 'bad' is the opposite of commendation, however one may call it. In everyday language we use positive value terms to commend things or persons with a certain respect. For example, healthy food is food that is commendable with respect to improving or maintaining one's health and a good action is one that is commendable 'all things considered'. From this point of view, competent speakers agree on this general component of the core meaning of value terms. If an honest and sincere speaker calls action A better than action B, all things considered, and still chooses action B or appraises B more than A, then

he lacks linguistic competence or makes a performative mistake. Even experts disagree, however, about what *really* makes an action commendable—and the proposed criteria in my view constitute part of the noumenal meaning of the respective use of ‘better than’.

As appealing as this position might seem, I am reluctant to identify the core meaning of ‘good’ and related positive value terms entirely with commendation within a given domain. Hare himself restricts this analysis to *moral* uses of ‘good’, which he considers entirely different from other uses,³⁰ and there may be everyday uses of evaluative adjectives whose core meaning goes beyond mere commendation. Many aspects of goodness hinge on factual matters and at least in relatively homogeneous speaker communities people may agree on many readings of value terms way more than one might suspect at first glance. For example, phrases like ‘a good hunting knife’ (instrumental goodness), ‘a good holiday trip’ (hedonic goodness), and even ‘a good life’ (ethical goodness) may have core meanings that go beyond mere commendation and noumenal meanings upon which even non-experts converge upon sincere reflection. Despite some disagreement on the details, there seems to be vast agreement on what constitutes a good hunting knife or makes a good holiday trip, and maybe also negative soft constraints take part of the core meaning: A knife with a lousy handle cannot be good, a trip on which both the hotel and the weather is horrible cannot be good, and a life of stress, unhappiness and misery can hardly be called a good life. Since there is also reason for doubt that moral and non-moral uses can be distinguished clearly, even seemingly moral uses might be based on a fixed set of fuzzy criteria within a speaker community. Is hedonic goodness a moral concept? For the Kantian, it is perhaps not, but for many classical utilitarians it is, and as long as we are not talking about what is *really* good in the hedonic sense, we may agree that being good in this sense means conforming to our subjective preferences and tastes, for instance, which vary from person to person. That being said, commendation always seems to be part of the core meaning of positive value terms, and their noumenal meaning may be understood as different attempts to find criteria as to why the actions, objects, or persons that satisfy the core meaning are commendable—and from such an attempt it might very well follow that, judging from the proposed noumenal meaning, they are not commendable at all. So DAT can accommodate Hare’s prescriptivist intuitions while at the same time allowing for the possibility that other philosophical positions about evaluative language may *also* be correct, by making the commendatory aspect of value terms part of their core meaning and explaining some cases of indirect value disagreement among competent speakers as disagreement about what the respective expressions really mean, viz., about their noumenal meaning.

³⁰ See Hare (1999: 140).

5.4. *Further Objections and Replies*

Despite its merits there are some objections to DAT that I will address and attempt to defuse in the following paragraphs.

A rather obvious objection is based on Hare's arguments and other variants of the Paradox of Analysis. In light of these arguments, does it even make sense to claim that value terms have a noumenal meaning? Would it not suffice to identify the evaluative meaning of value predicates as some sort of hidden imperative, as Hare suggests? In my opinion, this critique misses the point. Hare's position and that of many other noncognitivists may be described as the view that value terms are not associated with any fixed set of criteria into which they could be lexically decomposed, hence there might also be no defining characteristics that would justify talking about a noumenal meaning. Instead, these theories explain, in various competing ways, their core meaning by the ways these terms are connected with emotive responses, prescriptive uses, or their commendatory nature. But these are standpoints in moral philosophy, and from the perspective of DAT these are fine and compatible with the more modest aim of DAT to explain certain cases of value disagreement that do not seem to be based on direct, content-based disagreement. Expressed in terms of DAT, noncognitivists argue that the noumenal meaning of some or all value terms is (a) not given by descriptive features of the noumenon alone, because there are none or we do not have sufficient epistemic access to them, but by some features of the psychological or social role of those terms, or (b) that the noumenon of value terms consists of certain psychological or social facts that reflect and explain the commendable nature of their core meaning. This standpoint is compatible with DAT, and a major selling point for DAT as a semantic theory is precisely the fact that it can explain value disagreements like (7) between two speakers who look at the same issue (the noumenon) from vastly different metaethical stances.

Another, perhaps more serious objection concerns the semantics/pragmatics distinction and has also been raised from time to time against contextualist and relativist positions in the Philosophy of Language.³¹ The idea is to keep semantic representations minimal—though not 'crazy minimal' (Recanati 2006) like in Cappelen and Lepore (2004)—and resort to pragmatic explanations for the rest. For example, Väyrynen (2013) defends a pragmatic approach to thick evaluative concepts like 'lewd'. Would it not be better to bite the bullet and generally declare core meaning the 'real', primary meaning of expressions and abandon noumenal meaning altogether? This amounts to explaining disagreements like (1)–(8) in purely pragmatic terms, similar to what de Sá (2008) has suggested for the relativists' puzzles. It would go beyond the scope of this article to address this issue, especially since authors tend to presume their own semantics/pragmatics distinctions

³¹ See for example the critique by Bach (2005) on Recanati (2004).

and different answers can be given for different types of evaluative expressions. One powerful objection to a pragmatic approach, however, is worth mentioning in the current context: If noumenal meaning is generally explained in pragmatic terms, then this seems to amount to giving up truth-conditional semantics altogether, because the core meaning of many expressions is often semantically underdetermined and does not contribute enough to the truth-conditions of an utterance. Such a theory has undesirable consequences. Many expressions do make a precise truth-conditional contribution to the whole utterance in the form of noumenal meaning on which experts generally agree, even when experts do not completely agree or when their definitions are literally false or incomplete. For many purposes, the deficiencies of natural language use can be ignored, since any good theory is by its very nature highly idealized. To give an example, taking half-integer spin as a defining characteristic of fermions and integer spin as a defining characteristic of bosons might turn out to be inadequate in the future, but is this an argument against the adequacy of these definitions *now*? I believe not. Such revisions often occur in science, for example the Ur-meter was replaced by a definition based on the speed of light and the definition of a second was changed from one based on solar days to one based on the radiation of cesium 133 atoms. Despite such differences in the detail, we can say that both definitions sufficiently approximate the noumenon and are therefore adequate to a certain degree in many contexts of use. For example, a speaker who utters ‘This doorway is 2.20 meter high’ may be said to fully understand the truth-conditional contribution of ‘meter’ in her utterance *as long as* she does not use it in a way that does not sufficiently approximate a meter. To make this clear, Wittgenstein’s dictum of the beginning of this article is false, since mastery of the noumenal meaning is not required on behalf of speaker competence and core meaning may be too imprecise or unspecific to yield definite truth-conditions. But we may still say that a speaker understands an utterance in a given context if the truth-conditional core meaning sufficiently approximates the truth-conditional noumenal meaning of the expressions involved in accordance with the respective standards of precision that are in place in the given context. One might even go further and replace understanding with understanding to a certain degree, as long as it is kept in mind that for many purposes of coordinating behavior mastery of the noumenal meaning is not required. So there are two kinds of understanding in DAT. On the one hand, in successful communication a hearer may be said to fully understand an utterance without implicitly knowing its noumenal meaning. On the other hand, a speaker can be said to understand what a word or sentence in use really means when he or she grasps its noumenal meaning.³² Whenever experts mostly agree on it,

³² Neither of these types of understanding is based on speaker meaning or speech act content.

the noumenal meaning of an expression serves as a corrigendum of the core meaning; if noumenal meaning is discarded in this model, then the core meaning deteriorates into some vague notion of cognitive meaning suitable for many branches of linguistics but disconnected from truth-conditions. This is a high price to pay. Why give up truth-conditional semantics entirely if a truth-conditional DAT can explain, at the same time, the directedness of everyday language towards truth and reality, its use as a convenient tool for co-operation of behavior for which truth may sometimes only play a minor role, and the vagueness of truth-conditions of utterances in contexts of daily language use?

While I want to leave the exact nature of core meaning open for the time being, I have argued that it can be spelled out in terms of possibly semantically underdetermined truth-conditional meaning in a loose bundle view or the like, and noumenal meaning is truth-conditional just like Kaplan's character. This means that both of them are based on lexical decomposition of word meaning, which unfortunately also comes at a price. Any such account of word meaning must somehow defend itself against arguments that are directed against analyticity. Many of those arguments such as the Paradox of Analysis itself are not particularly convincing, and sufficient doubt has been cast on them elsewhere, so this section shall end with only some admittedly cursory remarks about these sort of criticisms which can be raised against many more semantic theories than just DAT. The idea behind them is generally that the lexical decomposition of a value term G into multiple criteria C_1, \dots, C_n or complex logical combinations of criteria is implausible as an account of lexical meaning, because it would make certain value judgments analytical and this fact would fly into the face of our intuitions. Note that without the last part there would be no problem and critiques rarely attempt to justify why our intuitions are worth a penny, but let us buy into this reasoning for the sake of the argument. One might reply to it that since Quine (1964) analyticity is often conceived as coming only to a certain degree rather than giving rise to analytic judgments with apodeictic certainty, and that DAT is compatible with this point of view because it states that (a) different speakers need only to associate similar core meanings with the same expression, (b) what the expression really means is often left to experts as part of a linguistic labor division, and (c) defining characteristics of both types of meanings may be false or otherwise not suitable, as long as there remains a way to identify the noumenon on the basis of those characteristics. For instance, although very unlikely, current chemistry might turn out to be incorrect and in need of revision because water turns out to be XYZ. Suppose it is XYZ. Then the noumenal meaning upon which experts unanimously agreed is inadequate, false, incoherent, or otherwise in need of revision due to changes in the underlying background theory. So analyticity is given the role that Quine reserved for it, not as a justification of infallible judgments, but as an explanation and indicator of

our willingness to give up certain claims about the meaning of expressions easier than others. Both meanings in DAT remain in principle revisable, and whether or not such a revision leads to a retrospective language change, attributing the new meaning to speakers prior to the change, may hinge on various linguistic and social factors. The resulting post-Quinean revisable and fallible notion of analyticity takes the edge off the Paradox of Analysis and related worries. Although it may have been part of the noumenal and core meanings of ‘atom’ that they cannot be split, making ‘Atoms are the smallest, indivisible building blocks of nature’ analytic in this weak sense, we all know that this definition eventually went up in flames.

There is a related critique following Moore (1903) based on a position that I wish to call ‘semantic primitivism’. In this view a use of ‘good’ points to some primitive property of goodness, ‘democracy’ means democracy or the property of being a democratic state, ‘electron’ stands for the property of being an electron, and so forth. In support of their theory, primitivists resort to variants of the Paradox of Analysis, Moore’s Open Question Argument, and sometimes an allegedly vicious definitory circle.³³

I do not have much to say about these kind of arguments except that I find them fairly unconvincing. They contradict existing knowledge and are pragmatically incoherent with existing linguistic practices. We already believe that a democratic state is somehow defined by some of its characteristics, and we already believe that a ‘good knife’ must have certain features and lack certain misfeatures. A primitive, non-decomposable concept theory also does not tell us anything useful about meaning and gives us no information about how to falsify statements containing the respective value expressions. That being said, it is worth emphasizing that DAT is compatible with the thesis that (a) certain referential expressions like proper names have no linguistically mandated core meaning—a speaker may associate a meaning with them, but need not do so by virtue of linguistic competence—, and the thesis that (b) the noumenal meaning of certain expressions such as natural kind terms is fallibly extracted from our investigations of the entities and ultimately fixed indexically. However, I hope to have made it plausible that value terms and predicates do not generally work this way.

³³ Moral intuitionists like Dancy (2004) seem to find the circle argument appealing. Since you have to justify the conceptual analysis of the conceptual analysis, and so forth, at some point moral intuitions have to kick in. I find the idea that you have to justify everything at any arbitrary level of semantic decomposition and the idea that intuitions could serve as a justification equally unappealing, and in any case the same allegedly vicious circle can be applied to any lexical analysis of any expression, not just moral terms. The bottom line of my reply is that from a Platonic point of view the circle is not vicious and there is no need to start all of your justifications with the Big Bang. Justification is a dialectic process that stops with agreement or in an aporetic stalemate.

6. *Summary and Conclusions*

Competent speakers associate a core meaning with value terms, and it has been argued that the commendatory nature of value terms may be the best candidate for this common denominator among speakers. Speakers also generally strive to match, upon further reflection, some fitting aspect of reality with the use of such terms. Even for syncategorematic expressions like ‘and’ various noumenal meanings have been suggested in the logical and semantic literature, and perhaps only few or none of them fully match the core meaning that may have evolved from our need to coordinate behavior by communicating with each other. When even logical expressions have controversial noumenal meanings, then it ought not come as a surprise that there are many different candidates for the noumenal meaning of value terms. Perhaps we tend to not be satisfied with the core meanings of value terms despite pervasive disagreement about suggestions for their noumenal meaning, because that is generally the way language works. Once we transcend our needs of communicating for some immediate needs, become curious about aspects of reality that go beyond a mere need of coordinating with others, and want to describe reality correctly, we naturally assume that there is something in reality, the noumenon, to which the use of a term corresponds—which does, of course, not imply that there really is such a thing. On the basis of this thesis I submit that dialogues like (1)–(8) do not exemplify mere negotiations but rather attempts to balance noumenal meaning and core meaning in a way that approximates the noumenal meaning to a correct description of the often ephemeral and purported noumenon. If the noumenal definition deviates too much from the core meaning, then we obtain a technical definition that may appear to be arbitrary. If on the other hand the noumenal definition deviates too much from the noumenon, as for instance new evidence is acquired, then it is revised, which may in turn trigger a slower revision of the core meaning. So after all, disputes about value terms that are not directly content-based may not be so different from disputes about theoretical terms or terms for abstract objects for which an ostensive definition will not do either. They are peculiar because of the widespread disagreement about their noumenal meaning, but this may be the result of their often rather complex multi-dimensional comparative structure and the complicated psychological and social phenomena with which they are connected. This approach is compatible with the metalinguistic negotiation thesis and should be understood as a *précis* and extension of it.

References

- Bach, K. 2005. "Context ex machina." In Z. G. Szabó (ed.). *Semantics versus Pragmatics*. Oxford: Oxford University Press: 16–44.
- Burge, T. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4:73–121.
- Burge, T. 1986. "Individualism and psychology." *Philosophical Review* 95: 345.
- Burgess, A. 2013. "Conceptual ethics i." *Philosophy Compass* 8 (12): 1091–1101.
- Burgess, A. and D. Plunkett 2013. "Conceptual ethics ii." *Philosophy Compass* 8 (12): 1102–1110.
- Cappelen, H. and E. Lepore 2004. *Insensitive Semantics*. Oxford: Blackwell.
- Chalmers, D. 2011. "Verbal disputes." *Philosophical Review* 120 (4): 515–566.
- Cresswell, M. J. 1990. *Entities and Indices*. Dordrecht: Kluwer.
- Cresswell, M. J. 1996. *Semantic Indexicality*. Dordrecht: Kluwer.
- Dancy, J. 2004. *Ethics without Principles*. Oxford: Clarendon Press.
- de Sá, D. L. 2008. "Presuppositions of commonality: An indexical relativist account of disagreement." In M. Garcia-Carpintero and M. Kölbel (eds.). *Relative Truth*. New York: Oxford University Press.
- de Sá, D. L. 2009. "Relativizing utterance-truth?" *Synthese* 170 (1): 1–5.
- de Sá, D. L. 2015. "Expressing disagreement." *Erkenntnis* 80: 153–165.
- Egan, A. 2010. "Disputing about taste." In R. Feldman and T. Warfield (eds.). *Disagreement*. New York: Oxford University Press: 247–286.
- Egan, A. 2014. "There's something funny about comedy: A case study in faultless disagreement." *Erkenntnis* 79: 73–100.
- Gärdenfors, P. 2000. *Conceptual Spaces*. Cambridge: MIT Press.
- Hare, R. 1952. *The Language of Morals*. Oxford: Oxford University Press.
- Hare, R. M. 1999. *The Language of Morals*. Oxford: Clarendon Press.
- Horgan, T. and M. Timmons 1992. "Troubles on moral twin earth: Moral queerness revived." *Synthese* 92: 221–260.
- Kaplan, D. 1989. "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals." In J. Almog, J. Perry, and H. Wettstein (eds.). *Themes from Kaplan*. Oxford, New York: Oxford University Press: 481–564.
- Kölbel, M. 2002. *Truth without Objectivity*. London: Routledge.
- Kripke, S. A. 1972. "Naming and Necessity." In G. Harman and D. Davidson (eds.). *Semantics of Natural Language*. Dordrecht, Boston: D. Reidel Publishing Co: 253–355.
- Laserson, P. 2005. "Context dependence, disagreement, and predicates of personal taste." *Linguistics and Philosophy* 28 (6): 643–686.
- Lewis, D. K. 1969. *Convention: A Philosophical Study*. Harvard: Harvard University Press.
- Lewis, D. K. 1980. "Index, Context and Content." In S. Kanger and S. Öhman (eds.). *Philosophy and Grammar*. Dordrecht, Boston: D. Reidel Publishing Co: 79–100.
- Ludlow, P. 2008. Cheap contextualism. *Philosophical Issues* 18: 104–129.
- MacFarlane, J. 2005. "Making sense of relative truth." *Proceedings of the Aristotelian Society* 105: 321–39.

- MacFarlane, J. 2012. "Relativism." In D. G. Fara and G. Russell (eds.). *The Routledge Companion to the Philosophy of Language*. New York: Routledge: 132–142
- MacFarlane, J. 2014. *Assessment Relativity: Relative Truth and its Applications*. Oxford: Clarendon Press.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.
- Marques, Teresa and García-Carpintero, M. (2014). "Disagreement about Taste: Commonality Presuppositions and Coordination." *Australasian Journal of Philosophy* 92 (4): 701–723.
- Meinong, A. 1971. "Über Gegenstandstheorie." In R. Haller (ed.). *Alexius Meinong Gesamtausgabe: Abhandlungen zur Erkenntnistheorie und Gegenstandstheorie*, Volume II. Graz, Austria: Akademische Druckund Verlagsanstalt Graz: 481–535.
- Moore, G. E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Penco, C. 1999. "Objective and cognitive context." In M. Benerecetti and F. Castellani (eds.). *Modeling and using context : Lecture Notes in Artificial Intelligence 1688*. Springer: 270–283.
- Plunkett, D. and T. Sundell 2013. "Disagreement and the semantics of normative and evaluative terms." *Philosophers' Imprint* 13 (23): 1–37.
- Plunkett, D. and T. Sundell 2014. "Antipositivist arguments from legal thought and talk: The metalinguistic response." In G. Hubb and D. Lind (eds.). *Pragmatism, Law, and Language*. London: Routledge: 56–75.
- Putnam, H. 1975a. "Is semantics possible?" In H. Putnam (ed.). *Mind, Language and Reality*. Cambridge: Cambridge University Press: 187–201.
- Putnam, H. 1975b. "The Meaning of 'Meaning'." In H. Putnam (ed.). *Mind, Language and Reality*. Cambridge, UK: Cambridge University Press: 215–271.
- Putnam, H. 1994. "The meaning of 'meaning'." In R. M. Harnish (ed.). *Basic Topics in the Philosophy of Language*. New York, London, Tokyo: Harvester Wheatsheaf: 221–274.
- Putnam, H. 2002. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge: Harvard University Press.
- Quine, W. V. O. 1964. *From a Logical Point of View*. Second revised edition. Oxford: Oxford University Press.
- Rast, E. H. 2009. "Context and interpretation." In J. M. Larrazabal and L. Zubeldia (eds.). *Meaning, Content, and Argument. Proceedings of the ILCLI International Workshop on Semantics, Pragmatics, and Rhetoric*. San Sebastian: University of the Basque Country Press: 515–534.
- Rast, E. H. 2011. "Classical possibilism and fictional objects." In F. Lihoreau (ed.). *Fiction in Philosophy*. Frankfurt: Ontos.
- Rast, E. H. 2014. "Context as assumptions." In F. Lihoreau and M. Rebuschi (eds.). *Knowledge, Context, Formalism*. Springer: 9–39.
- Recanati, F. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.
- Recanati, F. 2006. Crazy Minimalism. *Mind and Language* 21 (1): 21–30.
- Rosch, E. 1975. "Cognitive representations of semantic categories." *Journal of Experimental Psychology: General* 104 (3): 192–233.

- Rosch, E. 1983. "Prototype classification and logical classification: The two systems." In E. Scholnick (ed.). *New Trends in Conceptual Representation: Challenges to Piagets Theory?* Hillsdale: Lawrence Erlbaum Associates: 73–86.
- Stojanovic, I. 2015. "Evaluative adjectives and evaluative uses of ordinary adjectives." In D. Bekki and E. McCready (eds.). *Proceedings of LENLS12: Language Engineering and Natural Language Semantics*. The Japan Society for Artificial Intelligence.
- Sundell, T. 2011. "Disagreements about taste." *Philosophical Studies* 155 (2): 267–288.
- Väyrynen, P. 2013. *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*. New York: Oxford University Press.
- von Wright, G. H. 1963. *The Varieties of Goodness*. London: Routledge & K. Paul.
- Wittgenstein, L. 1969. *Tractatus logico-philosophicus*. London: Routledge.
- Zalta, E. 1983. *Abstract Objects*. Dordrecht, Boston, Lancaster: D. Reidel.