Speaker's reference, stipulation, and a dilemma for conceptual engineers

Max Deutsch[1]

Abstract Advocates of conceptual engineering as a method of philosophy face a dilemma: either they are ignorant of how conceptual engineering can be implemented, or else it is trivial to implement but of very little value, representing no new or especially fruitful method of philosophizing. Two key distinctions frame this dilemma and explain its two horns. First, the distinction between *speaker's* meaning and reference and *semantic* meaning and reference reveals a severe implementation problem for one construal of conceptual engineering. Second, the distinction between *stipulating meanings* and *conceptually analyzing* allows us to see why, on another construal of what conceptual engineering involves, the practice is neither a new nor neglected philosophical methodology. The article also argues that semantic externalism is not the root of the implementation problem for conceptual engineering, and that the usual rationale for adopting the practice, one that ties its value to the amelioration of "conceptual defects", is unsound.

Keywords Conceptual engineering · Speaker's reference · Stipulation · Semantic change · Conceptual analysis

✉ Max Deutsch
  maxdeutschhku@gmail.com

[1] University of Hong Kong, Pok Fu Lam, Hong Kong

1 Introduction

I will argue in what follows that philosophers who advocate *conceptual engineering* as a method of philosophy[1] fail to draw important distinctions, chief among these being the distinction between *speaker's* reference and meaning and *semantic* reference and meaning, and the related (as I will show) distinction between attempts to *stipulate* the meaning and reference of a term versus attempts to *reveal* or *characterize* a meaning or reference it already has. Once these distinctions are brought to bear on attempts to engineer concepts, it will be clear, I think, that conceptual engineering faces a serious *implementation problem*. Such attempts are efforts to stipulate new semantic meanings and referents for terms whose semantic meanings and referents are already fixed. However, although there is a class of important exceptions, stipulation usually does not have this power. For example, no single speaker, nor even large groups of speakers acting together, can simply stipulate that, henceforth, 'dog' shall mean *cats*. For the same reason, or so I will argue, one can't simply "engineer concepts" to suit one's purposes, no matter what, or how noble, these purposes might be.

I also seek to challenge the idea, promulgated by many of its recent defenders, that conceptual engineering represents a new (or at least neglected) and potentially revolutionary method of philosophizing, one capable of significantly supplementing, if not supplanting, older methods, such as the method of conceptual analysis.[2] I will argue that, given the sorts of things would-be conceptual engineers take to be examples of conceptual engineering, they face a dilemma: either we are ignorant of how conceptual engineering can be implemented, or it is straightforward to implement, but deeply uninteresting, involving no new technique, and ill-suited to solving, or even making genuine progress on, any philosophical problem.

2 The standard account of conceptual engineering

On what has become a standard account of the practice, conceptual engineering involves the recognition of "conceptual defects" combined with the attempt to repair these defects by improving or replacing the relevant concepts. For example, one interpretation of Haslanger's (2000) work on gender and race concepts is that Haslanger is not offering *descriptive* analyses of these concepts. Rather, she is offering *ameliorative* analyses of them. Our current gender and race concepts, such as 'woman' or 'Black' are, on this interpretation of Haslanger, defective, by failing

---

[1] These include Haslanger (2000), Cappelen (2018), Eklund (2014), Richard (2014), Nado (2019), Brun (2016), Chalmers (2011) and Burgess and Plunkett (2013a, b).

[2] Most present-day advocates of conceptual engineering trace their lineage to Carnap (1950) and the method of "explication". Most also draw a *contrast* between conceptual engineering/explication and conceptual analysis, and some are quite skeptical of the viability and fruitfulness of the latter. This is true of Nado (2019) and especially of Cappelen (2018), who has harsh things to say about philosophical work addressing the Gettier problem, for example, work that many regard as involving paradigmatic instances of conceptual analysis. (See Cappelen 2018, 26–27).

to be useful in the fight for social justice. So Haslanger proposes replacement concepts: Haslanger's analyses are analyses of the gender and race concepts we *ought* to have and use, not analyses of the gender and race concepts we *do* have and use.[3]

This *is/ought* distinction, drawn relative to concepts, is invoked by many other defenders of conceptual engineering. Here is Nado (2019) describing conceptual engineering and contrasting it with conceptual analysis:

> Conceptual engineers aim to improve or to replace rather than to analyse; to create rather than to discover. While conceptual analysts are interested in the concepts we *do* have, conceptual engineers are interested in the concepts we *ought* to have. Their project is prescriptive rather than descriptive. (Nado 2019, 3)

Eklund (2014) uses quite similar language but adds that conceptual engineering is a more *sophisticated* style of philosophizing, one not tied to our "folk" philosophical concepts:

> …while philosophers often have been concerned with our actual concepts or the properties or relations they stand for, philosophers should also be asking themselves whether these really are the best tools for understanding the relevant aspects of reality, and in many cases consider what preferable replacements might be. Philosophers should be engaged in *conceptual engineering*. Compare: when physicists study reality they do not hold on to the concepts of folk physics but use concepts better suited to their theoretical purposes. Why should things stand differently with what philosophers study? (Eklund 2014, 293)

Richard (2014), echoing some of Chalmers's (2011) view about the prevalence of verbal disputes in philosophy, claims that disputes over competing analyses of a concept can often be resolved only by asking the *normative* question of which concept we *should* employ (His example is 'free action'):

> Some philosophers tell us that to act freely would be to perform an act, the performance of which was not determined by conditions over which one has no control. Others tell us that to act freely is, roughly put, to perform an act such that one could have decided not to perform it (and would not have performed it, had one so decided). Yet other accounts are on offer…Why should we think that when we use the phrase 'free action' in speech or token it in thought, it is determinate that we are picking out the property isolated by one as opposed to another of these candidate analyses of free action?…it is not at all implausible that 'free action' does not determinately denote… if it does not, then all those interested in philosophical problems linked to the notion of

---

[3] Later in the main text, I discuss Haslanger's (2000) account of the concept 'woman' in particular. Still later, I discuss an alternative interpretation of this account, one given by Haslanger herself in Haslanger (2006) and again in Haslanger (2010) (though the alternative interpretation in Haslanger 2010 is applied primarily to her account of race concepts).

freedom can do is to describe the varying strands in our concept of free action and make recommendations, based on the interests we do or might have, as to how we might eliminate the vagueness of the concept. (Richard 2014, 401)

A last example of an advocate of conceptual engineering who adopts what I am calling the standard account of the practice is Cappelen (2018), who, after describing a variety of purported examples of conceptual engineering, describes their basic similarity very simply like so: "some concept is considered defective along some dimension, in some cases that deficiency can be ameliorated, and various proposals are made about how to best ameliorate". (Cappelen 2018, 33)

Proponents of this standard account of conceptual engineering tend to say very little about *what concepts are*.[4] I suspect this is because they don't think much *needs* to be said about this. Philosophers speak of concepts all the time; concepts are said to be the targets of conceptual analysis, for example, and, even among those skeptical of conceptual analysis, philosophy is said to consist to a significant extent in "conceptual clarification". Proponents of the standard account of conceptual engineering appear to think that there is no problem simply co-opting this "concept talk", trusting that philosophers, at least, will know how to interpret it, and so will know, in particular, what it is for a concept to be our "actual concept" '*F*', as opposed to the improved concept '*F*' we "ought to have and use". In fact, however, it is not at all clear what concepts are, and this makes it difficult to say even what it is that advocates of conceptual engineering are advocating, let alone whether there can be successful instances of the practice. Are concepts *mental representations that guide people's categorizations*, for example, as many psychologists of concepts would insist? Or is a concept instead the *meaning of a term*, one that "expresses" the concept, as perhaps philosophical conceptual analysts would insist? Is the "instead" in the last sentence even appropriate? Conceptual engineers need answers to these kinds of questions, if they want to say, as proponents of the standard account of conceptual engineering appear to, that conceptual engineers engineer *concepts*. Ok, but what *are* those? And what is it for a concept '*F*' to be a concept '*F*' we actually have and use, as opposed to the one we *should* be using?

3 *Speaker's* versus *semantic* meaning and reference

At this point, in order to define away some of the problem for the standard account of conceptual engineering described in the last section, I am going to regard conceptual engineers as subscribing to the view that concepts are—at least, if not only—the meanings of terms. Conceptual engineering is thus partly about terms in a language, such as English, entities that are syntactically individuated (I'll assume), but which come to have particular meanings and referents, ones they share with translations of these terms into other languages. So, as I am proposing it be

---

[4] Cappelen (2018) ends up rejecting the view that the targets of conceptual engineering are concepts, proposing instead that these targets are the intensions and extensions of terms. This proposal is similar to the one I will soon make in the main text.

understood, Haslanger's ameliorative analysis of the concept 'woman' is at least in part about the English word 'woman', and the issue of this word's meaning and reference. This gives us a way to interpret the idea, integral to the standard account of conceptual engineering, that conceptual engineers seek to improve our actual concepts: this involves taking some term, e.g. 'woman' or 'freedom', one that already has a meaning and reference (the actual concept the term expresses), and *changing* it, so that the term has a different—improved—meaning and reference (the concept the term should, and now does, if engineering has succeeded, express).

Although I think this understanding of conceptual engineering to some extent clarifies what conceptual engineers are up to, it raises several problems of its own. One of these is that there is no univocal sense of *the* meaning and reference of a term. Hence, there is no univocal sense of *the* concept a term expresses. This is because, while terms often do have meanings and referents that are best thought of as properties of the terms *themselves*—the terms' *semantic* meanings and referents—they can also be *used* by speakers to mean and refer to things other than their semantic meanings and referents. That is, the *speaker's* meaning and reference of a term can be different from its semantic meaning and reference.

It is difficult to overstate the theoretical significance of the distinction between *speaker's* and *semantic* meaning and reference.[5] But the *existence* of the two "kinds" of meaning and reference is clear from what we know about communication. We know, for example, that speakers can *use* a term, '*F*', to communicate about things that are not *F*s, i.e. to things not semantically referred to by (or in the semantic extension of) '*F*'. They can do this unwittingly, because they are ignorant or in error about which things are in fact *F*s, or intentionally, because they are being ironic, say, or because they want to mislead hearers.

The distinction is real, and theoretically significant, but why take it to pose any sort of problem for conceptual engineering? Roughly, this is because, although speakers are more or less free to use the terms of their language to *speaker*-mean and *speaker*-refer to whatever they like, they are not likewise free to determine the *semantics* of their language. In the next section, I argue that reflection on *linguistic stipulation* makes this problem for conceptual engineering more apparent.

4 Engineering as stipulation?

An interesting case of the use of '*F*' to intentionally speak of non-*F*s is *explicit stipulation* to the effect that one will use '*F*' to speak of non-*F*s. If I say to you, "for the next minute, I shall use 'dog' to mean *cats*", then for the next minute, as I say

---

[5] The distinction plays an important role in (a) Grice's (1989) theory of implicature and the related demise of "ordinary language philosophy", (b) the proper assessment of semantic contextualism, including epistemic contextualism, and the prospects for semantic minimalism (c) evaluating Radical Millianism about proper names, (d) determining whether the referential/attributive distinction, relative to uses of definite descriptions, is semantically significant, (e) understanding the significance of results in experimental philosophy, and (f) assessing theories of fictional discourse, metaphor, generic language, pejoratives, and slurs.

things like "dogs purr", "dogs don't bark", "there goes a dog", etc., you are right to interpret me as saying things about cats. It is reasonably clear, however, that I am nonetheless, and also, saying things about dogs for that minute. At the very least, the *words* I use say things about, and refer to, dogs. I am still speaking English for that minute, and, in English, 'dog' refers to dogs, not cats. *Qua* term of English, 'dog' has no cats in its extension.

One thing the case underscores is the need for the distinction described in the last section, a general distinction to describe the difference between what *I* am meaning and referring to with my uses of 'dog' for the relevant minute, and what the *word* 'dog' itself means and refers to, before, during, and after that minute. More significantly, the case illustrates that, while it seems possible for me to stipulate what I will *speaker* mean and refer to with my uses of 'dog', it also seems plain that my stipulation does not suffice to change the *semantic* meaning and reference of 'dog'—not even for a minute.

Consider a more extreme case. Suppose I publish a paper in which I declare that, henceforth, and for all time, 'dog' shall mean *cats*. What effect could this declaration have? Perhaps, afterwards, I and my followers (loyal and steadfast as they are) will now go around speaker-referring to cats with our uses of 'dog'. Clearly, though, neither my nor my followers' intentions or behavior will do much of anything to affect the semantic meaning and reference of the English term, 'dog'. Even if my explicit intention is to change the semantic meaning and reference of 'dog', it seems I am powerless to actually do so. Importantly, even if groups of other speakers start getting in on the act—my followers' followers, their followers, etc.— the only result will be a change in what certain groups of speakers now speaker-mean by 'dog' (along with a fair bit of miscommunication with others). 'Dog' will go on semantically meaning and referring to what it always has.

To prevent misunderstanding: I am *not* claiming that semantic change is impossible. Such changes do occur, even without anyone explicitly stipulating that they should, as various case studies in historical linguistics suggest. ('Meat', for example, has gone from meaning *food in general* to meaning *food that is the flesh of an animal*.) Rather, my claims are, first, that a speaker's stipulation that an existing term shall now have a different semantics doesn't suffice to *give* the term that new semantics. Second, we are, all of us, including would-be conceptual engineers, ignorant of what more must be done, or what else must occur, in order to bring about semantic changes. We don't know, for example, *how many* speakers, or *how often*, or for *how long,* these speakers must abide by an initial stipulation for any of this activity to effect a genuine semantic change. Indeed, it is not obvious, even, that speakers' *intentions* in *using* a term (which do suffice, I'll assume, for *speaker's* reference and meaning) ever suffice, by themselves, to effect genuine semantic changes. There is, in other words, an *implementation problem* for intentional semantic change: even if we are convinced, for whatever reason, that some existing term *ought* to have a different semantic meaning and reference, we don't know *what steps can be taken* to actually bring such a change about. Since even large groups of speakers stipulating the change and acting in accord with this stipulation won't *implement* the change, what *will* implement it? Conceptual engineers, as I am understanding them here, advocate making intentional changes in terms' semantic

meanings and referents. They therefore owe us an account of *how* they can actually succeed in doing so.

If the eventually proffered account were to claim that conceptual engineers can succeed by stipulative fiat—that all that is required for changing a term's semantic meaning and reference is someone stipulating that this change shall occur—then, clearly, there would be a problem: in many cases, including my imagined cases involving 'dog' described above, stipulation, by itself, is obviously not up to the task. In the imagined cases, I didn't—and couldn't by stipulation alone—change the semantic meaning and reference of 'dog'. So, I didn't—couldn't—successfully engineer the concept 'dog', not if the supposed mechanism of conceptual change is mere stipulation. This makes it questionable whether Haslanger has engineered, or can successfully engineer, the concept 'woman', or whether philosophers inspired by Richard (or Nado, Eklund, or Cappelen) can successfully engineer the concept 'free action'. Perhaps Haslanger has changed what she and some of her readers speaker-mean by 'woman', and perhaps a conceptual engineer might convince some speakers to speaker-refer to something other than the semantic reference of 'free action' when using the term. But this seems like a rather trivial and easy thing to do. Surely it is not the sort of thing the exciting terminology of "conceptual engineering" was designed to describe.

Of course, the complaint I am making is fair only if conceptual engineering really is the attempt bring about, by mere stipulation, changes in the semantic meanings and referents of terms. More than most other recent defenders of conceptual engineering, Nado (2019) suggests that this *is* what it is.[6] Although she speaks (unhelpfully, I think) of *concepts* as the targets of conceptual engineering, she also describes examples as instances of "pure stipulation" (14, 19) or "revisionist stipulation" (11). She does not elaborate this as I have; that is, as a matter of these stipulations being stipulations concerning the semantic meanings and referents of terms. But her account seems compatible with this elaboration, and many of her examples seem to be examples of attempts to stipulate new semantic meanings and referents for terms.[7]

---

[6] Haslanger (2000) is perhaps Nado's inspiration here. Haslanger, in expressing advocacy for ameliorative analyses of gender and race concepts, writes that "[i]n the limit case of an analytical [ameliorative] approach the concept in question is introduced by *stipulating* the meaning of a new term, and its content is determined entirely by the role it plays in the theory" (Haslanger 2000, 224; my emphasis). Note, however, that Haslanger is here describing stipulating the semantics of a *new* term. I discuss this kind of stipulation in the next section of the main text.

[7] Nado draws a distinction between *semantic* and *functional* conceptual engineering, one useful for understanding and assessing an objection to conceptual engineering due originally to Strawson (1963) to the effect that conceptual engineering *changes the subject* instead of shedding light on the subject it intends to address. Nado convincingly argues that this objection doesn't apply to functional conceptual engineering, whose aim is to preserve the function(s) of a concept, not its semantics. Despite this, functional engineering, even on Nado's account, proceeds via *semantic change*: the functional engineer tries to preserve function *by* changing meaning. So there is a sense in which functional engineering is no less semantic than semantic engineering. In both cases, the mechanism of conceptual change is a change in the semantics of a concept/term. This is why, despite Nado's *semantic/functional* distinction, I regard her view as compatible with, and perhaps even committed to, the view that conceptual engineering involves changes in the semantic meaning and reference of a term.

A case in point is Nado's discussion of a hypothetical conceptual engineer who proposes the following, bizarre "revisionist analysis" of free will:

Free will: x possesses free will iff x is an H_2O molecule. (Nado 2019, 6)

Interestingly, although Nado claims that "[n]ot even the staunchest revisionist would take this proposal seriously" (6), the reason she eventually gives is *not* that no one *could* engineer the concept 'free will' in the way the hypothetical proposal proposes. In fact, on Nado's view, this is perfectly possible, it is just that the result would be a concept of 'free will' that is "inefficient" for reasons I need not enter into here.[8] But *is* it really possible to engineer the concept 'free will' in the way proposed by Nado's hypothetical proposal?

Compare Nado's hypothetical conceptual engineer to someone who proposes the following, equally bizarre, "revisionist analysis" of dogs:

Dog: x is a dog iff x is a cat.

Could *this* proposal succeed? If the proposal is simply the stipulation that the semantic reference of the term 'dog' shall be cats, then it's difficult to imagine how it could.[9] As we have seen, attempts to stipulate new semantic meanings and referents for existing terms in a language are, in plenty of cases, doomed. Of course, the proposal could be intended merely as a recommendation about how to *use* 'dog': "use 'dog' to mean *cats*". But then, while possibly successful, the proposal is pretty deeply uninteresting. Sure, speakers can abide by the recommendation; they can *use* 'dog' to speaker-mean *cats*. But so what?

The same thing goes for Nado's hypothetical conceptual engineer attempting a radical "revisionist analysis" of free will: either it does not work, because one can't simply stipulate a new semantic meaning and reference for the term, 'free will', or it "works", but only because it encourages some speakers to speaker-mean that x is an H_2O molecule when they say "x possesses free will".

Although, as she points out, conceptual engineers would not take the hypothetical revision of the concept 'free will' Nado discusses seriously, I think the proposals

---

[8] On Nado's account, conceptual engineers seek "efficient" concepts to replace the defective ones with which they start. This is a matter of the new concepts being better suited to serving various functions and purposes. The issue of efficacy is mostly irrelevant to the point I am making in the main text, which can be put as follows. There are two questions for any Nado-inspired conceptual engineer. First, *can* the new concept replace the old by stipulation alone? And, second, is the replacement a more efficient concept than the concept it replaces? The argument I am making in the main text is that the answer to the first question is often "no". So, often, the question of whether an instance of attempted conceptual engineering is successful is answered in the negative, before the question of efficacy even arises.

[9] Taken at face value, the proposal is actually the obviously false proposal that nothing is a dog unless it is a cat, and vice versa. Clearly, that's not something one could stipulate to be so, nor does it concern concepts or the meanings of words. Note, though, that this is true of Nado's hypothetical proposal about free will too: it says something about free will and H_2O molecules, something no one could stipulate to be so, and something not about concepts or words. Despite this, in the main text, I am regarding the hypothetical proposal as a revisionary analysis of the *concept* 'free will', understanding concepts as (at least) the semantic meanings of terms. However, my view is that another problem with enthusiasm for revisionist analysis is that it is based on a deeply flawed picture of (non-revisionary) philosophical analysis, which is not a method of analyzing *concepts* at all. Philosophical analysts analyze knowledge, freedom, etc., *not* the concepts of these things.

they *do* take seriously suffer from the very same problem. Suppose, for example, that a Richard-inspired revisionist metaphysician were to offer the following as a stipulative analysis of the concept 'free action':

> Free action: *x* is a free action iff *x* is not determined by factors over which the agent has no control.

The question, as before, is whether this stipulation has any chance at all of succeeding. Imagine that, as a matter of fact, some actions determined by factors over which the agent has no control are in the semantic extension of the English term, 'free action'.[10] If this is so, and if, as it seems from the cases we have looked at so far, this is something that *can't* be altered simply by stipulative fiat, then it seems that we have here yet another case of an unsuccessful attempt at conceptual engineering. One could *try* to change the semantic meaning and reference of 'free action' simply by stipulating a new semantic meaning and reference for it, but this attempt seems doomed to fail. At best, the Richard-inspired metaphysician will convince some speakers to speaker-refer to all and only actions not determined by factors over which agents have no control when they *use* 'free action'. But 'free action' will go on semantically referring to some actions that *are* determined by factors over which agents have no control, just as it did, as we are supposing, before the attempt at conceptual engineering.

Or consider Haslanger's ameliorative analysis of the concept 'woman', which can be rendered, in brief, like so:

> Woman: *x* is a woman iff *x* is a person who faces subordination on the basis of perceived biological features indicating a female role in reproduction.[11]

Haslanger's analysis, conceived as the attempt to change, by stipulation, the semantics of the term 'woman', seems incapable of doing what it is intended to do. Suppose that the semantic extension of 'woman', in English, includes some people *not* subordinated on the basis of perceived biological features indicating a female role in reproduction. Haslanger's analysis hasn't changed this, if it's true, and it is difficult to see that it could. Once again, perhaps there are certain consequences of the proposal having to do with speaker's reference: maybe Haslanger and her followers now speaker-refer only to subordinated people with their uses of

---

[10] This is to imagine something that Richard perhaps rejects, namely that 'free action' "determinately denotes" some property distinct from the property described in the revisionist's stipulation. But this doesn't affect the point I am making, which is just that *if* 'free action' determinately denotes, then mere stipulation is powerless to change this. In any case, even if 'free action' fails to determinately denote, stipulation is also powerless to change *this* fact about the semantics of the term, or so I would argue.

[11] I have borrowed this abbreviated version of Haslanger's analysis from Nado (2019). See Haslanger (2000) for the full version.

[12] Again, I don't deny that a term's meaning and reference *can* change. The question is: What will turn a mere stipulation about the semantics of 'woman' into a genuine semantic shift? Haslanger can try to stipulate a new semantics for 'woman', and some of her readers can speaker-refer and speaker-mean in

'woman'. But, given that one can speaker-refer to pretty much anything with one's uses of 'woman', it is unclear why this should be regarded as any sort of advance.[12]

5 Stipulative introductions

So far, I have argued that if conceptual engineering is understood as the attempt to change, by mere stipulation, the semantic meaning and reference of a term, then, in many cases, such attempts won't succeed. One can stipulate that the *speaker's* meaning and reference of one's *use* of a term be different from its semantic meaning and reference, but it is not clear that this deserves to be called "conceptual engineering", or that this possibility represents a worthwhile, fruitful method of philosophizing. Attempts to effect genuine semantic changes in an existing term by mere stipulation seem bound to fail.

However, there is a class of important exceptions to the claim that the semantics of a term can't be simply stipulated as one pleases. These are cases in which one needs or wants a *new* term for an object, phenomenon, or property for which one's language does not yet have a term. I call these cases of "stipulative introduction". They are cases in which new terminology (in the sense of new *syntax*) is introduced and added to the language, along with a stipulation intended fix its semantics. Another case from Nado (2019) illustrates the sort of thing I have in mind. At one point, in order to show that conceptual engineers are "permitted to generate new concepts by pure stipulation", Nado herself "generates a new concept", writing "I hereby stipulate that a 'brollop' is a desk chair with five legs". (19) As she goes on to point out, once 'brollop' is introduced in this way, one can "generate truths" at will: "brollops are chairs", e.g. This appears to show that there are cases in which one *can* successfully determine the semantics of a term by a simple act of stipulation.

There is an obvious difference, however, between the brollop case and the cases described in the previous section. In the brollop case, there is a term, 'brollop', that has, as yet, no fixed semantics. Surely this is *why* Nado can successfully stipulate a semantics for it. But no matter how sincere or authoritative I am when I say "I hereby stipulate that a 'dog' is a cat, or that possessing 'free will' is being an H_2O molecule, or that a 'free action' is an action not determined by factors over which the agent has no control, or that a 'woman' is a subordinated person", it seems equally clear that I don't—can't—succeed in these cases. The relevant terms already have semantic meanings and referents, ones seemingly immune to "stipulative revision". In these cases, I do not manage to generate truths simply via my stipulative act. "Women are subordinated people", for example, isn't now magically true, post-stipulation. Nor is "dogs purr". Nor is it magically false that

---

[12] Again, I don't deny that a term's meaning and reference *can* change. The question is: What will turn a mere stipulation about the semantics of 'woman' into a genuine semantic shift? Haslanger can try to stipulate a new semantics for 'woman', and some of her readers can speaker-refer and speaker-mean in accordance with her stipulation. But without more details filled in, it is not at all clear that any of this activity will ever *implement* Haslanger's desired semantic change.

"there are free actions determined by factors not under the agent's control" or that "no H_2O molecule possesses free will". Nevertheless, the brollop case does show that new terms can be introduced with a semantics fixed by pure stipulation. The question is: What is the significance of such cases for conceptual engineering?

Somewhat surprisingly, it seems that some of what its defenders describe as "conceptual engineering" amounts to no more than the stipulative introduction of new terminology. I say "surprisingly" because there is nothing especially remarkable about stipulative introduction. It is certainly not some new, particularly fertile method of philosophizing, one that might supplement or supplant philosophical conceptual analysis. It takes no special insight or skill to stipulatively introduce a term, for example, and it seems utterly incapable of solving any genuine philosophical problem. Stipulative introduction won't reveal whether we have free will, all women are subordinated, or knowledge is justified true belief. In fact, it seems that most of its value derives from *syntactic convenience*: via stipulative introduction, we can replace longer descriptions ('desk chair with five legs') with a shorter, single term ('brollop').[13]

This is not to say that stipulative introduction has no place or value in philosophy. Sometimes we really do need new terms for things for which we lack conveniently short strings of letters or words. And not just in philosophy—in science, logic, mathematics, and everyday life, we sometimes have use for convenient names, or predicates consisting of one or a few words, terms with which we can replace our clunkier descriptions. At some stage in economic theory, for example, it became useful to introduce 'gross domestic product' to refer to a somewhat complex property possessed by national economies, one useful for understanding and comparing rates of economic growth. Economists *could* use a lengthy description every time they need to refer to this property, but 'gross domestic product' is more convenient, and 'GDP' even more so. There are countless cases like this in every theoretical context, cases in which something, or some property, important for theorizing, needs a convenient label. Stipulative introduction provides a straightforward way of introducing these labels and specifying their semantics. In other words, there are such things as "technical terms". Technical vocabulary often has its semantics fixed by stipulative introduction.

Like all other theoretical endeavors, philosophy has its share of technical terms introduced via stipulative introduction. Nado mentions 'supervenience' and 'credence'.[14] Others include 'grounding', 'intension', and 'secondary quality' along with many, many others. Do these cases deserve to be called cases of conceptual engineering? Describing the economist who introduced the technical term 'GDP' as having engaged in "economic conceptual engineering" seems like an overblown,

---

[13] I am open to the idea that, in some cases, perhaps especially in contexts of social/political speech, stipulative introductions do more than offer a convenient shorthand for lengthier descriptions. The introduction of the term "sexual harassment", for example, might well have been a remedy for what Fricker (2007) describes as "hermeneutical injustice". Still, I think these are special cases. Generally, stipulative introduction of new terminology is a remedy for inconvenience.

[14] Although the example of 'credence' is not exactly stipulative introduction, but instead an example of a superficially different variety of stipulation that I describe in the next section as "stipulative addition".

too-fancy description for what he did. This suggests that philosophers who introduce technical philosophical terms should not be credited with engaging in some newfangled, potentially revolutionary method of philosophy—they are just stipulatively introducing new terms, doing nothing much different from what Nado has done with 'brollop'.

Unlike Nado's brollop case, *useful* stipulative introductions are usually introductions of terms for phenomena theorists need or want to invoke for better theorizing. 'Supervenience' labels a relation important for theorizing about the mind–body problem, for example. And, of course, *identifying* this relation and *recognizing* its importance took some philosophical ingenuity. This is true about useful stipulative introductions generally: *labelling* the relevant property is rather trivial, identifying it and recognizing its significance is not. The economist who introduced 'GDP' insightfully recognized that GDP is useful for theorizing about economic growth.

Is this identification and recognition of theoretically significant phenomena what defenders of conceptual engineering are defending? Perhaps so. Consider, once more, Eklund's (2014) motivation for pursuing projects of conceptual engineering: like the theoretical physicist, relative to "folk" concepts of physics, philosophers, according to Eklund, need to think about whether our "actual" or "folk" philosophical concepts are the "best tools" for their "theoretical purposes".[15] Now, if this means only that philosophers should be thinking about whether the phenomena *for which we already have terms* exhaust the phenomena to which we might need to appeal in mature philosophical theory, then it is obviously good advice. *Of course* we should always be on the lookout for those objects and properties that might be useful for explaining whatever it is we seek to explain. And, in such cases, we can and should stipulatively introduce terms we can then use to conveniently refer to these entities and properties. The stipulative introduction of the relevant terminology is relatively trivial, Eklund and other defenders of conceptual engineering might admit. What is not trivial is the activity of identifying and recognizing phenomena not yet identified or recognized as important to sophisticated philosophical theorizing.

But did philosophers really need to be told any of this? Is Eklund's advice, as I am interpreting it, something that philosophers have not been heeding? Have defenders of conceptual engineering reminded philosophers of something they forgot, or perhaps never knew, namely that, sometimes, there is a need for technical terms in philosophy? I think it is obvious that the answer to these questions is "no". It is not as if philosophy proceeded entirely in terms drawn from ordinary language before Carnap (1950) came along and suggested that some technical vocabulary

---

[15] A similar motivation for conceptual engineering is described in Eklund (2015).

[16] Some advocates of conceptual engineering suggest that some of what gets *called* "conceptual analysis" in philosophy is actually, in at least some cases, *unwitting* conceptual engineering. But, since stipulating the semantics of a technical philosophical term is so *obviously* a different sort of undertaking than specifying the conditions on the application of an existing term, it is very hard to take this suggestion seriously. Philosophers know when they are simply stipulating meanings and when, by contrast, they are proposing analyses.

might be needed too. Rather, from its inception, philosophy has traded in technical terms, stipulatively introducing them as the need arises, just as every other theoretical enterprise has done.

Conceptual analysts, no less than any other sort of philosopher, sometimes need technical terms, and they show no aversion to using them in their analyses. Plenty of philosophers of mind use 'supervenience' in analyses of the relation between the mental and the physical, for example. So the *contrast* Eklund, Nado, and other defenders of conceptual engineering draw between conceptual analysis and conceptual engineering is puzzling: if conceptual engineering is or involves the stipulative introduction of technical terms, then conceptual analysts (and every other sort of philosopher) are already fully engaged in the practice. The recommendation to conceptually engineer instead of, or in addition to, conceptually analyzing thus offers no real guidance, and, frankly, seems a bit confused.[16]

6 Explication and stipulative addition

The argument to this point can be summarized as follows. Conceptual engineers face a dilemma. Conceptual engineering is either *stipulative revision*, the attempt to stipulate new semantic meanings and referents for terms whose semantic meanings and referents are already fixed, or else it is the *stipulative introduction* of technical terms. If it is the former, the prospects for success seem dim. Stipulative revisions might not be strictly *impossible*, but there is a severe-seeming implementation problem facing stipulative revision: How can the water of stipulation and speakers' intentions be transmuted into the wine of genuine semantic change? If conceptual engineering is instead the stipulative introduction of new technical vocabulary, then success is fairly trivially within reach. But there are plenty of cases of successful, useful stipulative introductions in philosophy already, including many made in connection with conceptual analysis. Conceptual engineering is neither a new nor neglected method of philosophy, if it amounts to the stipulative introduction of technical terms.

Does this dilemma for conceptual engineering overlook a third possibility? Perhaps conceptual engineering is neither stipulative revision nor stipulative introduction. Perhaps it is rather the *stipulative addition* of a new semantic meaning and referent to an existing term that already has one or more fixed semantic meanings and referents. Some cases of Carnapian "explication" seem intended as cases of stipulative addition, including some of Carnap's own examples. Carnap suggests that 'fish', for example, has been explicated to mean *piscis*, where 'piscis' is stipulated to have a more "exact" definition as 'cold-blooded aquatic vertebrate'. (Carnap 1950, 5–6) One understanding of this example, perhaps Carnap's own, is that 'fish' retains its "prescientific" meaning and extension, according to which

---

[17] I think insensitivity to the *speaker's/semantic* meaning distinction also creates confusion about examples like 'fish' and 'fruit'. Someone who says "don't put that tomato in the fruit salad, it is not a fruit!" says something *semantically* false. She might nonetheless communicate—speaker mean—something true.

whales allegedly count as fish, it is just that it is given, by stipulation, an additional meaning, a "scientific" one according to which whales are excluded. The picture is that 'fish' is now, post-explication, ambiguous: in some contexts (fish market?) its extension includes whales, in others (marine biologist's lab?) not.

A problem with this example, one it shares with many other alleged examples of explication, is that it assumes something like a *descriptivist* account of the semantic meaning and reference of "prescientific" terms. However, on an alternative view, applied to the example of 'fish', it was a *discovery* that "whales are not fish" is true, something it could not have been, had 'fish' not *already had* no whales in its semantic extension. Similarly, "tomatoes and olives are fruit" is true, *as we have discovered*. There's no sense in which "tomatoes are not fruit" is true, and it is not that now, thanks to scientific explicators, there is also a sense in which it is false. Rather, "tomatoes and olives are fruit" has been true all along, despite the existence of large numbers of speakers, now and in the past, who would deny it.[17] No one needed to do any explicating of 'fruit' in order for there to be a sense of 'fruit'—*the* sense of 'fruit'—according to which "tomatoes and olives are fruit" is true.

Putting problems with specific examples to the side, however, it seems clear that there is no general reason to deny that stipulative addition, and thus explication, if stipulative addition is the right model for it, can be implemented. For example, it appears possible to simply stipulate that one is an 'adult' just in case one has reached age 18. 'Adult' can then be used in its stipulated sense in some contexts, for some purposes (voter registration, e.g.), without this implying that the stipulated sense *replaces* its old sense. We can still say truly in some contexts, for example, that this or that 17-year-old is a borderline case of an adult, while also saying truly, in other contexts, that those same 17-year-olds are definitely not adults.[18] However, stipulative additions, while possible, are simply stipulative introductions, minus the introduction of new terminology. They are stipulative introductions of new meanings and referents for *old* terms instead of stipulative introductions of new meanings and referents for *new* terms. They are introductions of technical *senses*, not introductions of new technical *terms*. But, if this is right, then there is only a superficial difference between stipulative addition and stipulative introduction, and there is no genuine third possibility overlooked by the dilemma I have posed for conceptual engineers.

Philosophy, like other theoretical practices, does indeed sometimes attach technical senses to existing terms. Consider the use of 'credence' in epistemology. Presumably, the semantic meaning and reference of the ordinary English term 'credence' is quite different from the meaning and reference stipulated for it when it is invoked in epistemological theory. Is stipulating this meaning and reference for 'credence' an example of epistemological conceptual engineering? If so, then, again, there's nothing particularly new or neglected about conceptual engineering. Conceptual analysts do it all the time. Epistemologists, invoking their special sense of 'credence', use it to analyze disagreement and disagreement-based skepticism, for example. 'Credence', in the technical sense, labels an important property, one

---

useful for epistemological theorizing, a property epistemologists had good reason to identify and then use in constructing their theories. But none of this was unknown to them. No one had to insist that they stop analyzing and start engineering.

If the advice to conceptually engineer amounts to no more than, "good job, philosophers, keep doing what you're doing!", then fine. But the rhetoric surrounding recent discussions of the practice suggests that this is not the view. Many advocates of conceptual engineering are *complaining*: they are claiming that there is something *wrong* with philosophy, and especially with conceptual analysis. Cappelen (2018) combines his boosterism for conceptual engineering with skepticism about all "descriptive philosophy", including, presumably, descriptive conceptual analysis. And Nado (2019) prefaces her advocacy for conceptual engineering with the depressing speculation that there might not be a single, successful example of conceptual analysis anywhere in philosophy—no wonder, then, that she favors conceptual engineering as an alternative. But the belief that conceptual engineering might be a fruitful replacement for, or supplement to, a more traditional method, like conceptual analysis, is unjustified, if I am right that conceptual engineering is either (attempted) stipulative revision, or else stipulative introduction or addition. Stipulative revision faces an implementation problem that should make us doubtful of its prospects. And stipulative introductions and additions are common practice already; indeed they are already a *component* of many conceptual analyses. So, in neither case should dressing these practices up by describing them as "conceptual engineering" convince us that they represent a special antidote to what is allegedly ailing more traditional methods of philosophy.

7 Revelation, externalism, and the root of the problem

An irony in the antipathy some conceptual engineers feel towards conceptual analysis is that at least some of the projects nowadays billed as projects of conceptual engineering are better interpreted, *not* as attempts at conceptual engineering, but rather as something closer to good old-fashioned conceptual analysis. One interpretation of Haslanger's views, for example, takes her analyses of gender and race concepts as ameliorative analyses, instances of attempted conceptual engineering. This is how I was understanding Haslanger's views earlier. But that is just *one* interpretation of Haslanger, and it is not even the interpretation that Haslanger herself currently accepts. These days, Haslanger (2006) suggests that her analysis of 'woman', for example, might well be the correct *descriptive* analysis of our actual concept 'woman', not just the concept 'woman' we ought to have and use. Put in the way I prefer, on this alternative interpretation—the interpretation it seems Haslanger herself now accepts—her proposal concerns the semantic meaning and reference of the term 'woman', one to the effect that the semantic extension of this term includes all and only people subordinated on the basis of perceived biological features indicating a female role in reproduction. It is not merely the proposal that the semantic extension of 'woman' *should* be as just characterized. It

is the proposal that this is how its semantic extension *is* correctly characterized.[19] Thus, on this interpretation of Haslanger, there is no variety of stipulation that is in any way relevant to assessing her analysis. And, from the perspective taken in this paper, that is a good thing: attempted stipulative revision of the semantics of 'woman' can't clearly be implemented, and stipulative addition of a new meaning and reference for 'woman' is too easy, with no obvious benefit.

Haslanger (2006) has raised concerns about the *counterintuitiveness* of her proposals when taken as descriptive analyses: many speakers do or would intuit in ways that conflict with these analyses. Saul (2006) argues that this counterintuitiveness shows that we should regard Haslanger's analyses as instances of attempted conceptual engineering.[20] But the concern over counterintuitiveness greatly overestimates the evidential force of intuition. An analysis can be *correct*, even if it conflicts with powerful, widespread intuitions. Counterintuitiveness, all on its own, is therefore never a reason to reject an analysis.[21] Hence, I think that, despite its counterintuitiveness, we ought to take seriously the idea that Haslanger's analysis of 'woman' is *the* analysis of 'woman'—the correct, descriptive, philosophical account of *what it is* to be a woman.[22]

Indeed, philosophy is, and should be, in the business of overturning mistaken intuitions, platitudes, and theories about what counts as knowledge, freedom, goodness, responsibility, causation, justice—and being a woman. When it succeeds in this, philosophy is *revelatory*: it reveals something important, and perhaps quite surprising and counterintuitive, about the philosophical phenomena that are the subject matter of philosophy. Sometimes there is a need for technical terms; these

---

[19] Haslanger has long argued that gender is a "socially founded", but fully objective, feature of people. But, as I read her, this is supposed to be a *discovery* about the nature of gender. It is not simply that she has redefined 'woman' so that it now refers to a socially founded feature; it is rather that it does, right now, so refer, and has all along. Viewing her analyses as revelatory nicely explains this connection between Haslanger's social constructionism and the goal of her analyses. However, revelatory analysis can also potentially reveal that a feature one might regard as socially founded is not in fact socially founded. For example, such analysis might reveal that individual races are not socially founded features, and are, instead, something like *breeding populations*. For excellent critical discussion of both revelatory analysis (which he calls "revisionism") and the idea that races are breeding populations, see Glasgow's (2009) *A Theory of Race*.

[20] In Saul 2006 Saul suggests that there's a difference between an *intuition* that some perhaps merely hypothetical *x* is an *F* and a sincere *application*, by an ordinary speaker, of the *term* '*F*' to some actual *x*. Her complaint about Haslanger's analyses, construed as descriptive analyses, is that they conflict with ordinary applications (she calls this the "linguistic counterintuitiveness worry"), and she grants that their conflict with intuitions proves little. But it is no more plausible that the criteria ordinary speakers employ in applying the term 'woman', say, are the conditions on its *correct* application than it is to think that correct analyses can't conflict with widespread intuitions about hypothetical cases.

[21] Some philosophers, including an anonymous referee for this journal, are not as comfortable as I am with the possibility of true but deeply counterintuitive philosophical analyses. I should therefore emphasize that the dilemma for conceptual engineers posed in earlier sections does not depend on the views I am here defending concerning what I am calling "revelatory analysis". One can agree with me about the dilemma without agreeing with me about revelatory analysis, or, in particular, the interpretation of Haslanger as offering a revelatory analysis of womanhood.

[22] For the same reason, we ought to take seriously the idea that the JTB theory of knowledge is the correct theory of knowledge, as Weatherson (2003) and Hetherington (1999) have argued, despite the existence of near universal intuitions to the contrary.

can be useful tools for potentially revelatory philosophical analysis. But there is no reason to think that the introduction of technical vocabulary is a philosophical end in itself. It is merely one tool, useful for what ought to be the primary goal: accurately describing and understanding the philosophical phenomena we seek to describe and understand. Advocates of conceptual engineering are offering terrible advice, if they are advocating that philosophers simply fiddle with terms, stipulating new technical vocabulary, or new senses for older terminology. That kind of fiddling is not always entirely pointless, but it has a point, when it does, only in the service of some better end—for example, an improved, *and perhaps utterly revelatory*, understanding of knowledge, justice, or womanhood.

Their advice is even worse, if advocates of conceptual engineering are advocating that we attempt to stipulatively revise the semantic meanings and referents of our existing philosophical terms. As I have argued, stipulation by itself does not have that power, so it is unclear what conceptual engineers can do, post-stipulation, to implement their proposed semantic revisions. So far, my argument for this, supplemented by a few vague hints about the mechanism of semantic change, has been via cases: here is an example of a failed attempt at stipulative revision; here is another—and so on. Can something more be said? *Why* is it that we generally *can't* simply stipulatively revise semantic meanings and referents, replacing them with "better" ones? What is the root of the problem with conceptual engineering conceived as stipulative revision?

The distinction between *speaker's* and *semantic* meaning and reference provides part of the answer, as I have urged. It does seem possible, in many cases, to stipulate what one will *speaker* mean and refer to with one's use of a term. And this possibility can produce the illusion, if one fails to recognize the distinction between the two, that one is capable, via stipulative fiat, of changing the *semantics* of an existing term. But this leaves it unexplained why the semantics of an existing term is immune to stipulative revision in the first place.

A proposed explanation of this, one explored in some of the recent literature on conceptual engineering, appeals to *semantic externalism*, the view that the determination of semantic meaning and reference—the *metasemantics* of terms—is partly a matter of factors external to the beliefs, desires, intentions, and intentional behavior (e.g. stipulative acts) of speakers.[23] I am, myself, a convinced and committed externalist. In fact, I take externalism to be a prime example of revelatory philosophy. Work by Putnam, Burge, and Kripke has revealed that long-standing, entrenched views about meaning and reference were deeply mistaken, and philosophers now have a far superior—though purely descriptive!—picture of the metasemantic determinants of meaning and reference.

I do not, however, take externalism to be the root of the problem for stipulative revision. In essence, the reason is that *changes* in semantic meaning and reference are *compatible* with externalism. That is, the semantic meanings and referents of our terms can be externalistically determined, as I think they are, even if these meanings

---

[23] Cappelen 2018 (see especially Chapter 7) discusses the apparent incompatibility between externalism and conceptual engineering. This is also the topic of Koch (2018).

and referents can also *change*, as pretty much everyone assumes. By wide consensus, for example, the name, 'Madagascar', has gone from semantically referring to a portion of the African mainland to now semantically referring to the large island off Africa's east coast. But this does not show, as perhaps even Evans (1973) (the first to introduce the 'Madagascar' case) would agree, that *no* version of a Kripkean externalist metasemantics for names is correct. For even though earlier uses of 'Madagascar' trace back, causal-historically, to a portion of the African mainland, current uses trace back, causal-historically, to the island instead. A *re*-grounding of 'Madgascar' has occurred. Its semantic meaning and referent was, at some stage in the history of its use, determined anew, and differently. But the story of how this re-grounding occurred could perfectly well be an *externalist* story all the same. That is, we can still maintain that it is partly in virtue of externalist, causal-historical factors that the name's semantics *shifted*, from once referring to the mainland to now referring to the island.

Reflection on cases like the 'Madagascar' case shows that there is no easy inference from the fact that externalist metasemantic groundings—for example, Kripkean "initial baptisms"—*occurred in the past* to the conclusion that, since no one can act to change the past, no one can act to change the semantics of an existing term.[24] Terms can be *re*-grounded on new entities or phenomena. Existing terms, having been involved in an *initial* baptism, can be used in a *later* baptism—of something different. And this can, as it has in the 'Madagascar' case, lead to genuine semantic changes, changes that pose no threat to broadly externalist metasemantic views. But if semantic shifts can occur, consistently with externalism, then it seems mistaken to hold that externalism stands in the way of stipulative revision. Stipulative revisions can be viewed as attempts to *intentionally* bring about semantic shifts that *could* occur non-intentionally. However, since, in principle, externalism is no bar to the occurrence of the shifts in the non-intentional case, it seems unlikely that it would block them in the intentional one.

But the issue, as I have emphasized throughout, is not whether conceptual engineering, conceived as stipulative revision, is strictly impossible. The stipulative revision horn of the dilemma I have posed for conceptual engineers concerns *implementation*: even if the correct metasemantics for terms allows, as externalism does, that a term's existing semantic meaning and referent *can* change, there remains the question of how this change can be intentionally wrought. After all, externalism *does* imply that semantic shifts require corresponding shifts in external

---

[24] Cappelen (2018) claims that, on externalism, "[e]ven if we had all the information about the metasemantics of a term (about the use patterns, the histories, the sources of information, the interaction between the experts, etc.), it would appeal to factors that are in large part out of our control. For example, past facts play a role in determining the meaning of terms, but we can't change the past" (74). This suggests that, if externalism is true, intentional semantic changes would require the ability to change the past. I don't think Cappelen actually *believes* this suggestion—elsewhere, he says things that appear to contradict it. But the suggestion is there in the quoted passage.

factors: stipulation, by itself, never suffices. So, any would-be stipulative revisionist faces the implementation problem: She can *start* by attempting to stipulate revised meanings and referents, but what more can she do to actually *implement* these revisions?

The root of this problem—the implementation problem for conceptual engineering conceived as stipulative revision—is not externalism. On *any* metasemantic view that requires more, for an existing term to have a particular semantic meaning and reference, than just the *intention* on the part of some group of speakers to *use* the relevant term *as if* it had that very semantic meaning and reference, stipulative revisionists will face the implementation problem.[25]

Consider a version of descriptivism about names that requires that the *same* description be associated with a name by most of its users in order for it to be the meaning-giving and reference-fixing description for the name.[26] On such a view, large groups of speakers might nevertheless use the name as if it had a different semantics without it actually having, or coming to have, that different semantics. For example, a group of speakers could decide one day to try to stipulate that 'Einstein' is a synonym for 'the inventor of the atomic bomb'. On my own anti-descriptivist, pro-externalist view, this won't work, because the semantics of 'Einstein' has nothing to do with which or how many speakers associate which descriptions with 'Einstein'. But even if one does think this matters, one might nonetheless hold that it takes more to implement a revision of the meaning and reference of 'Einstein' than even large groups of speakers stipulating that it be so.

Examples of anti-externalist metasemantic theories which lead to an implementation problem for stipulative revision could be multiplied,[27] but this single example suffices to make the crucial point: the root of the implementation problem for conceptual engineering, conceived as stipulative revision, runs deeper than an apparent clash with externalist metasemantics. Externalism presents stipulative revisionists with the implementation problem, but so would plenty of anti-externalist metasemantic views. Again, on any view according to which more is required for the determination of semantic meaning and reference than speakers' intentions to *use* the relevant term in a stipulated sense, there will remain the question: What more must take place or be done in order for a semantic shift to

---

[25] I think of linguistic stipulations as simply intentions, usually made explicit in a stipulative act, to use a term in a particular way.

[26] Koch (2018) would classify this as an externalist view, given that he defines internalism as the view that the semantic meaning of a speaker's term is determined by that speaker's mental states and nothing more. (See Koch 2018, 3.) I doubt there are very many internalists when the view is defined this way. In any case, it is worth noting that certain forms of metasemantic descriptivism (even if not internalist on Koch's definition) are just as problematic for conceptual engineering as some anti-descriptivist views.

[27] Take a metasemantic view that holds that a term's semantic meaning and reference supervenes on *actual use*, where an actual use of a term is conceived as a speaker actually *applying* the term to this or that object or property. Since it's perfectly possible for speakers to apply a term in a way that fails to comport with any given *rule* for its application, including a stipulated rule to which all speakers agree, it is also possible, on such metasemantic views, for there to be a mismatch between a term's actual semantic meaning and reference and any stipulated semantics for it, even if every last speaker agrees to the stipulation and tries to abide by it.

actually be implemented? Not only do conceptual engineers rarely offer any answer to this question, a case can made that *there is no plausible answer to it*: we are, all of us, simply ignorant of the precise mechanisms of semantic change.[28] Here I agree with Cappelen (2018), who emphasizes what he calls the *inscrutability* of metasemantics: a variety of different factors, some "internal", some "external", matter to the determination of a term's semantic meaning and reference, but, in any given case, we don't know precisely which and we don't know precisely how. If metasemantics is inscrutable in this sense, then the intentional effort to render a semantic change in an existing term—an attempted stipulative revision—is just a shot in the dark. Stipulative revisionists can have no clear idea of *whether* or *how* or *when* their stipulations will render the relevant changes.[29] I think it is therefore obvious that the advice to *try* to stipulatively revise the semantics of our existing terms, advice that advocates of conceptual engineering are apparently offering, is terrible advice.[30]

8 Conclusion: "conceptual defects" and conceptual engineering's bad rationale

I have said very little thus far about why advocates of conceptual engineering regard the practice as worthwhile. I have avoided this because it is mostly irrelevant to my main argument. My dilemma for conceptual engineers implies either that it is obscure how to implement conceptual engineering, or else that it is trivially easy to implement, but that doing so buys little more than syntactic convenience. It is difficult to see how attempted conceptual engineering could actually *be* worthwhile,

---

[28] Koch (2018) argues that if something like Evans's (1973) view is correct, and our terms refer to the "dominant causal source" of the attitudes we express when using them, then we have what Koch calls "collective long-range control" over what our terms refer to. And we can exercise this control, Koch claims, simply by acting *as if* our terms have the referents we want them to have. Eventually, these acts can bring it about that the terms *do* refer to what we want them to refer to, by changing the causal source of the attitudes we express when using the terms. But the problem with Koch's picture, even granting an Evans-style metasemantics, is that there is no telling, in advance, that a plan to get large numbers of speakers to use t as if it refers to x will actually turn x into the dominant causal source of the attitudes speakers express when using t. People can make widespread *mistakes*, for example, using t even in relation to things that are not x's, despite the intention to use t only in relation to x's.

[29] Conceptual engineers might concede that it is unclear how to implement stipulative semantic revisions but deny that this matters. What matters, they might claim, is *using* the relevant terms in the stipulated way, to merely speaker-mean and speaker-refer in accord with the stipulation. This gives up on the idea that conceptual engineers seek semantic changes, but some of conceptual engineering's defenders (e.g. Nado 2019) seem eager to give up this idea anyway. The deeper problem with the concession is that it amounts to recommending that *we not aim to speak the literal truth when philosophizing using engineered terms*. I hope to elaborate this problem in future work. For a problem that I take to be similar, see Cappelen's (2018) discussion of reasons to avoid conceptual engineering merely for the sake of exploiting the "lexical effects" of a term (Cappelen 2018, 130–136).

[30] The most puzzling aspect of Cappelen (2018) is its insistence on the inscrutability of metasemantics and meaning change, on the one hand, and its advocacy of conceptual engineering, conceived as stipulative revision, on the other. To his credit, Cappelen is aware that this will strike many as puzzling. See Koch (2018) and Deutsch (2019) for criticisms of Cappelen's attempted resolution of this puzzle.

if this dilemma is correctly posed. Still, the issue deserves at least some comment here at the end, by way of conclusion.

Earlier, I claimed that the standard account of conceptual engineering views the practice as a remedy for "conceptual defects". Interpreted as a view about terms and their semantic meanings and referents, the idea is that our terms can and do have defective semantic meanings. This is how Cappelen (2018) for example, understands conceptual defects. And, according to Cappelen, conceptual engineering is the fix: we can remove these defects by changing the meanings of our terms. Conceptual engineering is worthwhile because it is the method by which we can repair and improve the defective semantics of our terms.

This account of the value of conceptual engineering clearly depends on the view that many of our terms, including many of our philosophical terms, such as 'knowledge', 'free action', and 'woman', are semantically defective. *Are* these terms semantically defective? Not if their purpose is to allow us to speak of, and communicate about, things like knowledge, free action, and women. A good way to speak of, and communicate about, knowledge, free action, and women is to use terms that semantically refer to these things, and the terms that semantically refer to these things include 'knowledge', 'free action', and 'woman'. So, the usual rationale for engaging in conceptual engineering is a bad rationale: since our terms are not, in fact, defective, relative to the purpose of using them to speak of their semantic referents, there is no need, and no value, in trying to improve them.

This brief argument against the usual rationale for conceptual engineering ignores the myriad ways in which, according to conceptual engineering's advocates, our terms are allegedly defective. Cappelen, for example, provides a long list of "representational defects", ones that conceptual engineers should be trying to repair.[31] I won't, here, go through these one by one because I think that Cappelen and others convinced of the usual rationale for conceptual engineering must be wrong about them. So long as our terms allow us to speak of, and communicate about, the things to which they refer—and all they need for this is to *have* semantic referents or extensions—then they are non-defective *enough* for our purposes, including, in the case of our philosophical terms, our philosophical purposes. Suppose, for example, that 'knowledge', like most other terms, is vague. (Vagueness is on Cappelen's list of alleged representational defects.) That doesn't prevent us from using it to speak of knowledge. The vagueness of 'green' doesn't prevent us from using it to speak of green things, after all. Maybe we can stipulate some precise terminology too, precise terminology useful for theorizing in epistemology. (Recall the case of 'credence'.) But if part of what we want to do is theorize *about knowledge*, then 'knowledge' is a pretty useful tool, semantically referring, as it does, to the very phenomenon about which we seek to theorize.

What about conceptual/representational defects allegedly inherent in terms for phenomena we regard as socially or politically significant (not that 'knowledge' is not one of these!)—'woman', 'Black', 'marriage', and 'immigrant', for example? Here too the rationale for conceptual engineering—repairing these allegedly

---

[31] See Cappelen (2018, 34).

defective terms—is misconceived. First of all, revelatory (but purely descriptive) philosophy can potentially reveal that these terms have a semantics that is surprising and counterintuitive—and this revelation, not any conceptual engineering, might then contribute to furthering the morally admirable goals (e.g. anti-sexism and anti-racism) for which philosophers who have discussed these examples aim. But, secondly, it is difficult, even in these cases, to make sense of the idea that the relevant terms are genuinely *semantically* defective. 'Marriage' and 'immigrant', for example, are non-defective enough to allow us to use them to speak of marriage and immigrants, or so it certainly seems. Of course, people have all sorts of false, and in some cases immoral, views about marriage and immigrants, including, to mention just two, that same-sex couples can't be married, or that the USA should ban all Muslim immigrants. But we need terms with the very semantic meanings and referents possessed by 'marriage' and 'immigrant' to speak of what these views concern, and to explain their falsity and immorality.

So, while the dilemma for conceptual engineering presented in earlier sections shows, I think, that conceptual engineering is either difficult to implement or else trivial, it also appears that it lacks a sound rationale. The usual rationale for the practice ties its value to the amelioration of "conceptual defects". I have argued that these defects are not really defects at all, on the assumption that the purpose of our terms is to speak of the phenomena to which they semantically refer.

If not conceptual engineering, what? The answer I favor is: philosophical argument, analysis, explanation, description, and theorizing. Conceptual engineering, with the exception of the introduction of useful technical terminology, doesn't belong in the philosopher's tool box. Then again: there are plenty of useful tools in there already.

References

Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis, 81*(6), 1211–1241.

Burgess, A., & Plunkett, D. (2013a). Conceptual ethics I. *Philosophy Compass, 8*(12), 1091–1101.

Burgess, A., & Plunkett, D. (2013b). Conceptual Ethics II. *Philosophy Compass, 8*(12), 1102–1110.

Cappelen, H. (2018). *Fixing language: An essay on conceptual engineering.* Oxford: Oxford University Press.

Carnap, R. (1950). *Logical foundations of probability.* Chicago: University of Chicago Press.

Chalmers, D. (2011). Verbal disputes. *Philosophical Review, 120*(4), 515–566.

Deutsch, M. (2019). Fixing language: An essay on conceptual engineering by Herman Cappelen. *Analysis, 79*(3), 574–578.

Eklund, M. (2014). Replacing truth? In A. Burgess & B. Sherman (Eds.), *Metasemantics: New essays on the foundations of meaning* (pp. 293–310). Oxford: Oxford University Press.

Eklund, M. (2015). Intuitions, conceptual engineering, and conceptual fixed points. In C. Daly (Ed.), *The Palgrave handbook of philosophical methods* (pp. 363–385). London: Palgrave Macmillan.

Evans, G. (1973). The causal theory of names. *Proceedings of the Aristotelian Society Supplementary, 47,* 187–225.

Fricker, M. (2007). *Epistemic injustice*. Oxford: Oxford University Press.

Glasgow, J. (2009). *A theory of race*. New York and London: Routledge.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Haslanger, S. (2000). Gender and race: (What) are they? (What) do we want them to be? *Noûs, 34*(1), 31–55.

Haslanger, S. (2006). What good are our intuitions? Philosophical analysis and social kinds. *Proceedings of the Aristotelian Society Supplementary, 80*(1), 89–118.

Haslanger, S. (2010). Language, politics and 'the folk': Looking for 'the meaning' of 'race'. *The Monist, 93*(2), 169–187.

Hetherington, S. (1999). Knowing failably. *Journal of Philosophy, 96*(11), 565–587.

Koch, S. (2018). The externalist challenge to conceptual engineering. *Synthese*. https://doi.org/10.1007/s11229-018-02007-6.

Nado, J. (2019). Conceptual engineering, truth, and efficacy. *Synthese*. https://doi.org/10.1007/s11229-019-02096-x.

Richard, M. (2014). Analysis, concepts, and intuitions. *Analytic Philosophy, 55*(4), 394–406.

Saul, J. (2006). Gender and race. *Proceedings of the Aristotelian Society Supplementary, 80*(1), 119–143.

Strawson, P. (1963). Carnap's views on conceptual systems versus natural languages in analytic philosophy. In P. A. Schilpp (Ed.), *The philosophy of Rudolph Carnap* (pp. 503–518). Chicago: Open Court.

Weatherson, B. (2003). What good are counterexamples? *Philosophical Studies, 115,* 1–31.