Please do not cite without permission

*non-ideal conceptual engineering*

Which Framework is Best Suited for Social Justice Theorizing?

**Introduction**

Conceptual engineering, as I will understand it, is a form of theorizing that proposes improvements to a linguistic practice. This construal is meant to be neutral with respect to alternative metasemantic frameworks. I make no commitment to the nature of concepts, or whether concepts are the relevant subject matter of conceptual engineering. Further, I make no commitment to the nature of linguistic meaning, or whether meaning is determined by linguistic practice.

I do, however, commit myself to the idea that conceptual engineering is about *change.* Specifically, change to a linguistic practice. And the ‘change’ in question is a matter of causal factors triggering a sequence of temporal events beginning with a linguistic practice with certain properties at t1, and ending with a revised linguistic practice, with different properties, at t2. In this sense, change to a linguistic practice is an *evolution*. It happens over time. The issue with this is that at any event in this temporal sequence, a linguistic practice, undergoing changes, can be pushed in different directions. The engineer, however, intends for the practice to be pushed in one direction, rather than another – specifically, an *improvement.*

Many conceptual engineering projects, in recent time, have proposed improvements to our linguistic practices with the aim of advancing pursuits in the project of social justice. I have in mind strategies such as Sally Haslanger’s (2000) ameliorative definitions of ‘race’ and ‘gender’, Katharine Jenkins’s (2016) trans inclusive account of ‘woman’, Kate Manne’s (2018) revisionary analysis of ‘misogyny’, and Robin Dembroff’s (2016) reappraisal of ‘sexual orientation’. Other examples abound.[[1]](#footnote-1) While each account offers a proposed improvement to our existing linguistic practices, certain questions can be raised. We can look *backwards* and ask: How plausible are such suggestions? And, we can look *forward* and ask: Should engineers continue to submit proposals? Such questions, however, can only be understood within a particular framework of conceptual engineering. There is no theory-independent vantage point of assessing the plausibility of an ameliorative strategy.

With this in mind, this paper introduces the notion of *non-ideal* conceptual engineering; a notion that matches, in very broad strokes, one in moral and political philosophy. In particular, *I will attempt to identify the most suitable framework of conceptual engineering for social justice theorizing*. And, the means by which I will do this is by assessing three competing frameworks of conceptual engineering along axes of *feasibility* – a standard against which normative theories are evaluated in non-ideal theorizing. This includes: (i) control, (ii) looping effects, (iii) right kind of reasons, and (iv) risk of perversion.

What makes this project non-ideal? It puts the quest for social justice, and real-world facts about oppression, first. It makes the choice between competing frameworks of conceptual engineering one that concerns advancements toward the resolution or amelioration of unjust structures. And, this is not such a bad means of decision-making. After all, there is little agreement amongst conceptual engineers – no one has settled the question of which framework of conceptual engineering is ‘right’. The considerations I explore in this paper might be able to tip the balance in favour of one framework.

Now, this does not mean that whatever framework comes out best is true. It simply means that it is the most suitable framework for social justice theorizing. The goal is to make sense of the existing ameliorative projects that have been offered, and whether we should continue to make proposals – after all, without a plausible story of how improvements to a linguistic practice can be implemented, it is unclear how useful ameliorative projects are in the fight against injustice. However, if there is independent reason to reject the truth of the theory that comes out best, or there are clear reasons that outweigh social justice, then it will turn out that *no* framework of conceptual engineering is suited for social justice purposes – even if one is ‘best’. Given this, my aim is not to argue that one theory stands out as best all-things-considered. This paper is a project, relative to the available facts, that examines which account of conceptual engineering can be put in the service of social justice.

This project differs, then, from another way we might construe non-ideal conceptual engineering. In this other sense, non-ideal conceptual engineering will take for granted a particular framework, and assess the value of a proposed improvement to a linguistic practice against actual and possible scenarios, rather than idealised ones. For example, we might assess whether an amelioration of our gender concepts, within a framework of conceptual engineering that aims at changing object-level facts, is worth pursuing given entrenched sexist social structures. However, this way of understanding non-ideal conceptual engineering requires finding a framework in which assessments of particular proposals can be made. This means identifying a theory that is able to tell us when and which ameliorative projects are able to be implemented out in the wild.

A note before I continue. The reader will notice an absence of *internalist* theories of conceptual engineering. By ‘internalist’, I don’t mean to limit myself to theories of meaning, but instead theories of conceptual engineering that discuss higher-order cognition more broadly. It is only recently that such theories have been introduced into the fold (Pollock 2020, Fischer 2020, Isaac 2020, and to some extent Podosky 2018). The reason for this omission is only due to limiting my discussion to frameworks of conceptual engineering that have been more prominently examined. It may turn out that internalist theories are better suited to social justice theorizing. However, I have reservations about this given that much injustice is material.[[2]](#footnote-2)

**1. Feasibility**

The notion of *feasibility* is most prominently explored in theorizing about the constraints that govern the designing of normative principles in moral and political philosophy. Importantly, the questions centre on whether normative theorizing must be constrained by what one can reasonably be expected to achieve. ‘Ideal’ theorists tend to reject this need on principles of justice; ‘non-ideal’ theorists accept such constraints (Valentini 2012, p. 4).

In recent literature, some theorists of conceptual engineering have, more or less explicitly, been interested in questions of *feasibility* *(*e.g., Podosky 2018, Fischer 2020, Andow forthcoming).[[3]](#footnote-3) Like moral and political philosophy, the questions centre on whether normative theorizing about a linguistic practice should be constrained by reasonable expectation. After all, if conceptual engineering is impossible, then it is not unreasonable for one to judge that it is pointless. Feasibility, as a constraint on conceptual engineering, ‘should be understood as the worry that the success rate of conceptual engineering may be low enough to make conceptual engineering not worth bothering with’ (Andow forthcoming). And, whether conceptual engineering is possible, or plausible, is sensitive to a number of different factors.

I will understand feasibility, as that which governs normative theorizing about a linguistic practice, to concern at least four axes of constraint: (i) control, (ii) looping effects, (iii) right kind of reasons, and (iv) risk of perversion. I will briefly explain each in turn.

*1.1. Control*

Part of the challenge of conceptual engineering concerns our ability to successfully implement an ameliorative strategy – to see a designed linguistic practice out in the wild. I needn’t commit myself to any set of success conditions. Whatever these conditions are, there must be some sensitivity to our ability to *control* the character and direction of a linguistic practice towards a proposed improvement. That is, there must some means by which a designed linguistic practice, developed by an engineer, can be intentionally propagated within a community of thinkers and speakers. Thus, the problem of ‘control’ refers to the extent to which the propagation of a *linguistic practice* is possible. This formulation is neutral with respect to which linguistic facts we must have control over (e.g., extension, usage, etc.)

*1.2. Looping Effects*

My interest is in the resolution of unjust social relations. Sometimes injustice can only be resolved through the amendment, and even creation, of social categories (Dembroff forthcoming). For example, if the social category of *woman* is unjust insofar as it excludes trans women, then justice requires amending this category to be more inclusive. Part of what is required to achieve this is the introduction of new linguistic practices, such as categorizing and application dispositions, which construct, or reconstruct, categories through *looping effects* (Hacking 1999, Mallon 2016).This is the mutual causal feedback between our linguistic practices (i.e., classification) and social kinds. As Robin Dembroff puts it, ‘by developing classification practices, we create social kinds, which in turn impact classification practices, and so on’ (forthcoming, p. 7). Thus, where ‘control’ concerns the propagation of a linguistic practice, this axis of feasibility concerns the chances of successful looping effects, brought about with an ameliorative strategy, that create or amend *social categories* (Podosky 2018, p. 13).

*1.3. Right Kind of Reasons*

Conceptual engineering projects are often motivated by particular reasons. For instance, a theorist might design a new linguistic practice surrounding ‘marriage’ with the aim that this practice will go some way towards resolving certain injustices, such as combatting discrimination against same-sex couples (Pollock 2020). However, it is entirely possible that the designed linguistic practice, if it ever were to be taken up, is adopted by members of a linguistic community for reasons other than those that motivated the engineer. After all, a linguistic practice might be propagated by preference for conformity, or a desire to guarantee efficient communication, etc. Thus, this axis of feasibility concerns whether conceptual engineers should be committed to the idea that their proposed improvement to a linguistic practice must be adopted for the *right kind of reasons*.

*1.4. Risk of Perversion*

Teresa Marques (forthcoming) has argued that conceptual engineering projects are permissible only when they are not *perversions*. ‘That means that they are [revisions] that (a) don’t have harmful consequences, and that (b) do not misapply a word to something unfitting the abstract values that the use of the word presupposes’ (forthcoming, p. 21). If this is right, then conceptual engineers must assess their proposal relative to the chances of being subject to *perversion*, rather than just the potential for amelioration. This axis of feasibility is closely related to control. If we don’t have control over a linguistic practice, then of course this means our proposal is at risk of perversion. However, it is important to assess the risk of perversion within conceptual engineering frameworks in which control is possible, albeit difficult.

*1.5. Feasibility and Social Justice*

One might ask: Why these axes of feasibility? To my knowledge, these are the concerns that have been raised by those interested in the limits of conceptual engineering. However, there is more to be said. Each axis of feasibility stands as a potential barrier to change. If we cannot *control* linguistic practices, then there is no point in trying to change them for the purposes of social justice; if we cannot amend or create social categories through *looping effects* with an ameliorative strategy, then there is no point in creating such strategies for the purposes of social justice; if we cannot expect people to adopt a linguistic practice for the *right kind of reason*s, then there is no point in proposing such a requirement for the purposes of social justice; and if we cannot develop a linguistic practice without some guarantee that it won’t be *perverted*, then we shouldn’t develop such a practice for the purposes of social justice. The question, then, is which framework of conceptual engineering will be able to meet the conditions of feasibility set out above?

**2. Frameworks**

Now that I have the axes of feasibility of the table, I want to introduce three competing frameworks of conceptual engineering that, to my understanding, are the most prominent in the literature: (i) The Worldly Account, (ii) The Alignment Account, and (iii) The Functional Account. After this, I will assess the relative merits of each theory against each axis of feasibility.

*2.1. The Worldly Account*

An early suggestion as to what would count as a successful conceptual engineering project was offered by Herman Cappelen (2018). He argued that the limits of conceptual engineering should be understood within a moderate externalist metasemantic framework, in the tradition of Kripke, Putnam, Burge, and Williamson. According to externalist theories of this kind, a range of factors ‘outside of the head’, such as past patterns of usage, communicative chains, and experts within a community, play a role in fixing reference.

Further to this, Cappelen’s understanding of conceptual engineering is strongly *metaphysical.* The success of conceptual engineering depends on actual changes to the world – the goal of ameliorative projects is to change reference, its intension and extension (2018, p. 138).[[4]](#footnote-4) Put differently, conceptual engineering is not about changing a linguistic practice *per se*, but changing a linguistic practice with effect of changing object-level facts. For instance, an account that aims to ameliorate linguistic practices with ‘woman’ is ultimately an account that aims to change what women *are.*

*2.2. The Alignment Account*

In response to Cappelen, Sarah Sawyer (2020, forthcoming) has argued that the ‘worldly’ construal of conceptual engineering, one that focuses on changing object-level facts, isn’t the best option to take within an externalist framework. Part of the reason for this is that Cappelen cannot account for *topic continuity*. Conceptual engineering must not change the relevant subject matter of inquiry, for this would not solve the problems that were of initial interest to us (Strawson 1963).

Some have understood the problem of topic continuity as a matter of whether conceptual engineering can preserve the extension of a term. If the extension of a term has changed, then the topic has changed – or so it goes. Sawyer disagrees with this. She does not think that the meaning of a *term* fixes the relevant subject matter. Instead, it is the reference of a *concept*, expressed by a term, that connects the term to a topic. Thus, what makes Sawyer’s account distinctive is a broadly Burgean distinction between the *meaning* of a term and the *concept* it expresses.

For Sawyer, the meaning or intension of a non-indexical term is externally determined – it is determined by social practices; stable patterns of use and deference within a community. Put differently, the meaning of a term is a characterization of the relevant subject matter that members of a linguistic community would settle on were they to reach reflective equilibrium (2020, p. 383). And, because patterns of usage change throughout time, the intension of a term, and therefore its extension, can change too. The question, then, is how extensions can change without a change in topic?

The answer: concepts, that are expressed by terms, fix the relevant subject matter. Like the meaning of terms, according to Sawyer, concepts are externally determined. However, concepts, in contrast to terms, are constituent elements of thought that representationally connect thinkers to a topic. And, the relationship between concept and the world is referentially direct: fundamental, non-conceptual/descriptive relations to objective properties.

What’s important to note is the relationship between concepts and term meanings. The concept expressed by a term fixes the relevant subject matter, and the meaning of a term is determined by a community’s characterization of this subject matter. Thus, the extension of a term can change and yet pose no problem for the continuity of a topic. What mustn’t change, in order to preserve topic continuity, is the extension of a concept.

On Sawyer’s account, then, conceptual engineering is the project of alignment. Specifically, it is about changing a linguistic practice in a way that aligns the extension of a term with the relevant subject matter. In Sawyer’s words,

If a revisionary analysis if correct and accepted, the effect is to bring the extension of the linguistic meaning of a term in line with the extension of the concept it expresses (i.e., in line with the relevant subject matter); it moves the linguistic practice closer to truth (2020, p. 391).

From this, we can see that Sawyer’s take on conceptual engineering isn’t metaphysical – it isn’t strictly about changing object-level facts. Instead, it is *epistemic*.[[5]](#footnote-5) It is about the accuracy of a community’s characterization or theory of a particular subject matter fixed by a concept.

*2.3. The Functional Account*

In contrast to both Cappelen and Sawyer, some have argued that conceptual engineering is ultimately about the relationship between function and linguistic practices. There are different ways of construing what this means.

*2.3.1. Functional Roles*

One way to think about function is to consider the goal or purpose behind a linguistic practice within a social and representational milieu. In the language of concepts, this understanding is common. Friedrich Steinle argues that ‘[o]ne of the fundamental characteristics of concepts is their directedness towards a specific goal’ (2012, p. 105); Ingo Brigandt (2010), claims that concepts have ‘epistemic goals’ that figure in explanations and inferences in scientific theorizing; P.F. Strawson (1963) talks about the ‘purpose’ of a concept, which is echoed by Frank Jackson (2011); and Manuel Vargas (2013) suggests that concepts do particular ‘work’.[[6]](#footnote-6) Moreover, Sally Haslanger takes a similar line in her understanding of what constitutes conceptual engineering:

The task is not simply to explicate the normal concept of X; nor is it to discover what things we normally take to fall under the concept we have in common; instead we ask what purpose is served in having the concept X, whether this purpose is well-conceived and what concept (or concepts) would serve our well conceived purposes(s)... (1999, p. 352).

Taking stock, Michael Prinzing argues that philosophical theorizing about concepts has a common theme:

All of these philosophers seem to have converged... on the same thought. The driving idea is that there are ‘intentions’ or ‘goals’ behind concepts; they have ‘jobs’; they are supposed to do certain ‘work’; they have a ‘point’, ‘purpose’, ‘role’, or ‘function’ (2017, p. 14).

Thus, on this way of thinking, a linguistic practice is engineered when there are changes to our use of a term, but the goal or purpose behind such usage is preserved (Haslanger 2020).[[7]](#footnote-7) This guarantees topic continuity.

*2.3.2. Etiological Functions*

Another way to think about function, one inspired by prominent theories in the philosophy of biology, is to consider the *etiological* or *proper* function of a linguistic practice (or concept). Such functions are those that exist in virtue of being beneficial to our ancestors. For instance, the function of a heart, which is the reason why it exists today, is explained by the fact that it pumped blood in our ancestors, contributing to their survival. Etiological functions are those that exist as a result of natural selection over generations.

This kind of view, offered as one kind of conceptual engineering, is endorsed by Mona Simion and Simon Kelp (2020). Concepts, at least some of them, continue to exist and operate in our social and representational milieu in virtue of being beneficial to our ancestors; they are selected-for given environmental pressures. Such concepts have a successful history of positive feedback in a thinking and speaking community. Thus, according to Simion and Kelp, one way of implementing an ameliorative strategy, which is under our control, involves designing an enticing linguistic practice that people will be disposed to copy. A concept is successfully engineered when its designed function becomes a function that explains its continued existence (2020, p. 12).

*2.3.3. Why Not Both?*

Given the foregoing, how should we construe the ‘functional’ account? It’s important to note that these accounts of function are not inconsistent. Sometimes we might be interested in changing a linguistic practice while *preserving* its goal or purpose (or trashing a linguistic practice in virtue of having a bad goal or purpose). Other times, we might be interested in designing a function for a linguistic practice such that, through selective pressures, it becomes a proper or etiological function. For instance, we might want a designed linguistic practice for ‘marriage’ to be selected-for, and at the same time want linguistic practices with ‘food’ to change while preserving its extant goal or purpose. Thus, I will treat the ‘functional’ account of conceptual engineering as inclusive of these two understandings.

An important note is that on the inclusive account of function that I am working with, there is no connection to meaning or concept identity (Nado 2019, cf. Fisher 2015, Prinzing 2018). It is the continuation of a function that must be preserved post-engineering, rather than semantic or topic continuity. The upshot is that conceptual engineering projects can be much more ambitious (Nado 2019).

**3. The Relative Merits**

I’m up to the point where I can assess the relative merits of each framework of conceptual engineering against each axis of feasibility. I’ll begin with the ‘worldly’ account.

*3.1. The Worldly Account and Feasibility*

When it comes to *control*, Cappelen is the first to admit that his account isn’t promising. The chances of implementing an ameliorative strategy seems impossible. After all, what fixes reference, the target of engineering on Cappelen’s account, occurs outside of the head. Because of this, our ability to know the reference of our terms is undermined. Therefore, so is our ability to control reference. In other words, according to Cappelen, ‘[t]he process governing particular changes [in reference] is typically incomprehensible and inscrutable’ (2018, p. 53). He recognizes that epistemic agents within a linguistic community are not fully, or perhaps not even mostly, informed of the relevant empirical facts that determine reference. And, to compound this problem, we might think that power-structures obscure the channels by which we might gain such knowledge. For example, we might be mistaken about who counts as an ‘expert’ in virtue of the fact that people occupy social positions that signal epistemic competence and trustworthiness, when in fact they are *not* competent nor trustworthy (Fricker 1999). Thus, any attempt to actively bring about changes in the reference of our words is bound to be highly unreliable.[[8]](#footnote-8)

Perhaps a benefit of Cappelen’s account, however, is that it does focus on changing reference, which allows for the possibility of *looping effects*. That is, through changes to our linguistic practices, we can affect which social categories exist in the world. This accommodates social justice projects that aim at amending or creating social kinds for the purposes of resolving unjust social relations (i.e., trans inclusive feminism, marriage equality, etc.). However, without a means of deliberately bringing about changes to object-level facts, it is unclear how much of a benefit to Cappelen’s view this really is.

What about the *right kind of reasons*? It doesn’t seem to me that Cappelen’s account requires any kind of motivation to adopt a linguistic practice. This might make the chances of changing a linguistic practice, for the better, much more likely. After all, control over linguistic practices, and therefore object-level facts, is hard enough without requiring that people take up a practice for specific reasons. However, if we think, like Joey Pollock (2020), that some reasons are *constitutive* of certain ameliorative projects, such as the reasons that should govern linguistic practices with ‘marriage’, then this is yet another stumbling block for Cappelen’s account.

Cappelen is also aware that in virtue of lacking control over semantic facts, conceptual engineering, if attempted to be put into action, runs the risk of *perversion*:

The fact that conceptual engineering is inscrutable and out of our control means that it is also possible (sometimes I think even likely) that those who try to achieve good ends through conceptual engineering will end up causing harms they didn’t intend. We have no prima facie reason to think the process is typically one that leads to amelioration rather than degeneration (2018, p. 159).

Of course, we should expect that bad things might happen when we try to implement an ameliorative strategy. However, this expectation must be weighed against our chances of success. On Cappelen’s account, we have no idea what the chances of success are, and thus all we are left with is the risk of perversion.

*3.2. The Alignment Account and Feasibility*

It seems that Cappelen’s framework of conceptual engineering is not suited to social justice theorizing. What about the ‘alignment’ account?

It’s hard to know how much better Sawyer’s framework of conceptual engineering is with respect to control. In one sense, it is better than Cappelen’s insofar as the goal is not to change object-level facts. This is one less step required in the process of conceptual engineering. Instead, conceptual engineering is a matter of alignment – the extension of a term must come to align with the relevant subject matter fixed by a concept that the term expresses. However, it is unclear how much control we have in this process of alignment. This is for at least two reasons.

Like Cappelen, Sawyer’s metasemantic framework makes it hard to know the reference of our terms. As such, it faces similar problems. It may be too much for us to fully grasp all of the linguistic activity that plays a role in determining meaning, such as past patterns of usage, or who the relevant experts are, and this undermines our ability to control semantic facts. The benefit of Sawyer’s account, however, is that even if we do manage to change a linguistic practice, in whatever direction, if the extension changes then this will not disrupt topic continuity.

There is a deeper issue with Sawyer’s account. Given that the relationship between concept, that which fixes the relevant subject matter, and the world is referentially direct, then we are left without a means of *knowing* or *reliably believing* when the meaning of a term has aligned with the reference of a concept. So, even if we grant that members of a linguistic community have epistemic access to the determinants of meaning, this does not entail that we are in any position to understand whether an intended change to meaning aligns with the relevant subject matter – the goal of conceptual engineering. This is a problem of ‘control’ not simply in the sense that we are unable to reliably manufacture change to semantic facts. Instead, it is a problem insofar as we do not have control over semantic facts *with the aim of aligning such facts with the extension of a concept.* We are unable to know when intended changes to the extension of a term aligns with the extension of a concept that it expresses.

Sawyer’s framework doesn’t do well on the axis of control. Perhaps it does better on the axis of looping effects? Unfortunately not. Remember, for Sawyer, it is the world, connected to a term via a concept, that guarantees topic continuity. And, in order to maintain the legitimacy of a conceptual engineering project, we must not change the topic. Thus, any change to the realm of (social) facts runs the risk of changing the relevant subject matter that an engineering project is supposed to be about. However, it is the changing of social facts that some conceptual engineers aim at in their attempts to achieve social justice.

For instance, suppose that a conceptual engineer is interested in changing social facts about what women are, or who counts as a woman. As such, she proposes an account that is trans inclusive, with the aim of revising linguistic practices such that an amended category of *woman* is ‘looped’ into existence.[[9]](#footnote-9) Apart from saying that it is ‘intuitively false’ that conceptual engineering changes the nature of women (or any other properties), Sawyer argues that such proposals transgress the limits of conceptual engineering (forthcoming, p. 10). After all, if there have been changes to object-level facts that connect a term to a subject matter, then the subject matter has changed – there is no single topic that persists through changes to the extension of a term. Thus, we have not preserved topic continuity. This is recognized by Sawyer, who says:

If the process of conceptual engineering changes, for example, what belief is, and does so, as Cappelen says, by changing the conditions that have to be satisfied in order for an object to fall into the extension of the term ‘belief’, then the property of being a belief is not a stable objective property that exists independently of the conditions that we associate with the meaning of the term ‘belief’ (forthcoming, p. 10).

Thus, looping effects, which change object-level facts, are not possible on the ‘alignment’ account. This goes against many conceptual engineering projects that aim at changing social, extralinguistic facts.

Does Sawyer’s account require that an improved linguistic practice be taken up for the right kind of reasons? This doesn’t seem necessary. If the goal is for the extension of a term, determined by linguistic practices, to align with the relevant subject matter, then this can be achieved, it seems, with people adopting a proposed linguistic practice for different reasons. For example, suppose that the relevant subject matter represented by the concept food excludes meat. The meaning of ‘food’ can be changed to match this subject matter through the adoption of revised linguistic practices with ‘food’ for different reasons – e.g., health, moral, environmental, fashion, etc.

Finally, Sawyer’s ‘alignment’ account makes it difficult to assess the risk of perversion. Of course, this is due to being ‘out of our control’ in the sense that we cannot reliably manufacture semantic change. But it also runs the risk of perversion insofar as the relevant subject matter is not epistemically available to members of a linguistic community. So there will always be some doubt as to whether an ameliorative project has been successful, which might mean we under- or over-shoot our target, or completely miss it altogether. Moreover, to give a more situated analysis, if the power to ‘decide’ whether change has occurred, as a matter of who decides whether we have gotten at the relevant subject matter, is distributed unjustly across a community, then perversion is a genuine concern. For example, suppose that attempts to redefine ‘woman’ fall short of full trans inclusion, and stop at those who have had sex-reassignment surgery. While the amelioration of a linguistic practice has occurred to some extent, it does not go far enough. This might be due to prominent voices in the community who advocate for this more narrow ameliorative project, or because the government has passed legislation about who counts as a woman. The problem is that if the goal is epistemic, as Sawyer’s account suggests, then the success of conceptual engineering will have to come up against the power that obscures our channels to knowledge and understanding.

*3.3. The Functional Account and Feasibility*

The ‘worldly’ account and the ‘alignment’ account appear to be inadequate frameworks for social justice theorizing. Can the ‘functional’ account do any better?

With the shift from meaning to function,[[10]](#footnote-10) having control over the process of conceptual engineering becomes much more plausible. After all, the goal is not to change a linguistic practice in order to bring about meaning change or preserve concept identity. With respect to social justice theorizing, the goal of conceptual engineering on the ‘functional’ account will be to change a linguistic practice that will ameliorate social facts. On the different ways of understanding what ‘function’ means, we might want to change a linguistic practice, but preserve its underlying goal or purpose; or else we might want to introduce a new linguistic practice, motivated by certain reasons, that will be selected-for.

Thus, controlling changes to a linguistic practice is, in the first place, much easier on the ‘functional’ account insofar as we are not nearly as epistemically limited. We do not need to know all of the linguistic activity that plays a role in meaning. In some cases, all we need to identify is the dominant linguistic practice in a context, and reveal the relevant goal or purpose of this practice via some process of rational interpretation of what we trying to achieve (e.g., Edward Craig-style (1990) conceptual ethnography). Sally Haslanger (2020) captures this in her example of our use of ‘family’.

According to Haslanger, the function of our concept of family is to organize our lives together with respect to the ‘coordination of domestic life, for example intimacy, sex, raising of children, economic partnerships, intergenerational transfers of traditions and property’ (2020, p. 251). In Western contexts, the typical arrangement that stabilizes coordination in domestic life consists of a husband, a wife, and biological offspring. The issue with this, however, is that the function of family thought and talk is contingently caught up in hetero- and bio-normativity. But this does not exhaust all domestic arrangements that can constitute a family. For instance, there are same-sex couples, adopted children, single or unmarried parents, and extended families. Acceptance of such domestic arrangements have been a result a societal pressure in the form of activism and material changes (2020, p. 251).

What this example shows is that the function of certain linguistic practices are epistemically available to us, and in virtue of this, such practices are able to be scrutinize and revised. It has been recognized, for some time now, that our existing practices with ‘family’ has been unjustly exclusive of other forms of domestic arrangement that can preserve the goal or purpose of such language.

Moreover, one of the issues with Cappelen’s and Sawyer’s account is that there are no clear-cut clues as to when meaning has changed, or when meaning has aligned with the relevant subject matter. But on the ‘functional’ account, there are obvious signs. We need only ask: Have social facts changed? Have trans and gender queer parents been accepted *qua* parents? Thus, not only is it possible for an engineer to recognize a linguistic practice, it is also possible for an engineer to make an informed judgement as to whether their project has been, at least somewhat, successful.

So far, I have only spoken about control with respect to one notion of function. What about control within an etiological function framework? Simion and Kelp (2020) have argued that this is possible. To reiterate, the mark of successful conceptual engineering is when a designed function a linguistic practice (concept) becomes a function that explains the existence of the practice (concept). The designed function must catch on and be stabilized within a community of thinkers and speakers. And, for Simion and Kelp, a conceptual engineer can achieve this by acting on the environment that will drive adaptive change. The environment needs to be tinkered with in order to drive demand for the designed function. Thus, provided that we can change the environment in this way, then we have some control over the direction and character of our linguistic practices.

How does the ‘functional’ account do on the axis of looping effects? It is evident in the foregoing example that looping effects are not only possible but available for us to understand and actively set in motion. After all, once upon a time it might have been inconceivable that our linguistic practices with ‘family’ might include application dispositions to same-sex couples. However, with revised classificatory practices, off the back of social movements, same-sex couples now fall comfortably, but not without resistance, within the extension of ‘family’. This example shows that the *amendment* of social kinds, through looping effects, is possible on the ‘functional’ account – but what about the *creation* of social kinds? This is also possible. Suppose that the goal of gender thought and talk is to keep track of people’s self-identification, or the internal maps that they use to navigate social reality. This goal has been preserved with the introduction of linguistic practices with ‘non-binary’. And, given how looping effects work, the classification of people as ‘non-binary’ has brought into existence a new social kind – non- binary people.

The feasibility of the ‘functional’ account also does well on the axis of right kind of reasons – though, it does better or worse depending on which notion of function one is working with. With respect to the notion of function that concerns the preservation of the goal or purpose of a linguistic practice, there doesn’t appear to be a necessary requirement that the revised practice be adopted for the right kind of reasons. For instance, suppose that the goal or purpose of practices with ‘food’ is to organize members of a community around a set of eating resources. And, suppose that an engineer, motivated by moral reasons, wants to revise practices with ‘food’ such that it no longer applies to non-human animals. While the engineer might like to see their product adopted for the reasons it was designed, it seems that they should be happy with their product being adopted for a range of different reasons – so long as such reasons are consistent with revised practice. Thus, the conceptual engineering of ‘food’ is successful even if people simply want to cut meat out of their diet for health reasons, or to save the environment, or to follow the latest trend.

In contrast to this, however, on the etiological story of function, the goal is not *merely* to change a linguistic practice, but for the practice to be adopted for the reason it was designed. That is, in the language of concepts, when an ameliorative strategy is ‘launched on a competitive market of concepts’, the hope is that the designed concept will ‘*be used to do what [it is] designed to do, and [it] will continue to be used in virtue of the fact that the way of thinking about the world [it] made available was beneficial to users*’ (Simion and Kelp 2020, p. 6 my emphasis)*.* Thus, the aim is not just to change linguistic practices, but for this change to be brought about and sustained in a particular way. One natural interpretation of this, at least when it comes to social justice concerns, is that a linguistic practice must be adopted for the right kind of reasons. Joey Pollock (2020) argues for this way of thinking when it comes to concepts for social equality, such as marriage:

...in ameliorative projects of the sort under consideration, engineers are motivated by the pursuit of moral goods – they are interested in contributing to the dismantling of oppressive social structures, institutions and systems of belief, and replacing them with those that will promote and sustain social equality. When, in implementing an ameliorated concept, engineers seek to transmit their own motivating reasons to the individuals who adopt the new concept, it is because acceptance of these reasons is partly constitutive of the sorts of changes that they wish to effect – individual who adopt the new concept ought (if the project is considered to be successful) to be appropriately motivated by specific moral considerations (2020, p. 88).

Pollock believes that reasons must be *constitutive* of certain changes. The ameliorative strategy is not solely about the adoption of a concept, but includes, as a matter constitution, a certain set of motivating reasons.

There are good reasons to think that some changes to linguistic practices must be adopted for the right kind of reasons. Being motivated by the right kind of reasons is often required to ‘overwrite specific existing problematic beliefs and dispositions with those that will contribute to promoting and sustaining social equality for the relevant group’ (2020, p. 90). However, it may turn out that this requirement is simply infeasible. It would be wonderful if everyone stopped eating meat because they were motivated by moral concerns. But, it seems that we should be happy with different motivations *so long as pernicious linguistic practices, and thus oppressive social facts, change.* Given that the world is unjust, in which many people are suffering as a result of being situated in systems of dominance and subordination, adopting a revised linguistic practice for the wrong reasons may be *good enough*. It will go a long way toward resolving oppressive social relations, even if it doesn’t go *all* the way – at least, not immediately.

A response might be that moral motivations are more robust. And if we want change, we must want stable change. Otherwise we risk falling back into old habits. However, I think we should still be happy with the adoption of a revised linguistic practice for different reasons, yet *continue* our efforts to convince people to accept better reasons.

Another response is that if we allow, for the success of conceptual engineering, that a revised linguistic practice is adopted for the wrong reasons, or mere conformity, or by accident owing to something else being selected-for, then this runs the risk of perversion. For example, if it turns out that the dominant reason why people have adopted revised linguistic practices with ‘food’ is that it will help solve climate change, then with the resolution of climate change, linguistic practices might revert back to applying ‘food’ to non-human animals. Clearly it is a good thing if climate change were to be resolved. However, there are harmful consequences for non-human animals if people do not accept that we ought to revise linguistic practices with ‘food’ to end the suffering of farmed animals.

**4. Taking Stock**

So, who wins? The foregoing reasons lean clearly in favor of the ‘functional’ account. Neither the ‘worldly’ account nor the ‘alignment’ account seem to do well with respect to control, looping effects, and the risk of perversion. In contrast, the ‘functional’ account does well with respect to all axes, however when it comes to the right kind of reasons, it might turn out that etiological functions are simply too hard to manufacture – at least for morally motivated linguistic practices.

This assessment of the relative merits of each framework of conceptual engineering does not settle the dispute as to which framework is correct. If it turns out that there is no way of plausibly defending the idea of the function of a linguistic practice or concept, then of course the ‘functional’ account should be abandoned. After all, the best framework of conceptual engineering for social justice theorizing is, minimally, one that exists. However, given that it is by no means settled which framework of conceptual engineering is correct, I take it that this means it is available to us to make a decision as to which framework we want to adopt on non-epistemic grounds. That is, the decision can be based on the goal of resolving unjust social relations.

**References**

Andow, James. Forthcoming. Conceptual Engineering is Unlikely to Work. So What? *Inquiry.*

Barnes, Elizabeth. 2016. *The Minority Body: A Theory of Disability*. Oxford: Oxford University Press.

Barnes, Elizabeth. 2019. Gender and Gender Terms. *Nous*.

Brigandt, Ingo. 2010. The Epistemic Goal of Concept: Accounting for the Rationality of Semantic Change and Variation. *Synthese* 177: 19 – 40.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering.* Oxford: Oxford University Press.

Craig, Edward. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Clarendon Press.

Dembroff, Robin. 2016. What is Sexual Orientation? *Philosophers’ Imprint* 16.

Dembroff, Robin. Forthcoming a. Real Talk on the Metaphysics of Gender. *Philosophical Topics.*

Fischer, Eugen. 2020. Conceptual Control: On the Feasibility of Conceptual Engineering. *Inquiry* DOI: 10.1080/0020174x.2020.1773309

Fisher, J.C. 2015. Pragmatic Experimental Philosophy. *Philosophical Psychology* 28 (3), pp. 412 – 433.

Fricker, Miranda. 1998. Rational Authority and Social Power: Towards a Truly Social Epistemology. *Proceedings of the Aristotelian Society* 98 (2): 159 – 177.

Hacking, Ian. 1999. *The Social Construction of What?* MA: Harvard University Press.

Haslanger, Sally. 1999. What Knowledge Is and What It Ought to Be: Feminist Values and Normative Epistemology. *Philosophical Perspectives* 13: 31 – 55.

Haslanger, Sally. 2000. Gender and Race: (What) Are They? (What) Do We Want Them to Be? *Nous* 34 (1): 31 – 55.

Haslanger, Sally. 2020. Going On, Not the Same Way in *Conceptual Engineering and Conceptual Ethics*, Alexis Burgess, Herman Cappelen, and David Plunkett (eds.), pp. 230 – 260. Oxford: Oxford University Press.

Isaac, Manuel Gustavo. How to Conceptual Engineer Conceptual Engineering. *Inquiry* DOI: https://doi.org/10.1080/0020174X.2020.1719881

Jackson, Frank. 2011. On Gettier Holdouts. *Mind & Language* 26 (4): 468 – 481.

Jenkins, Katharine. 2016. Amelioration and Inclusion: Gender Identity and the Concept of Woman. *Ethics* 126 (2): 394 – 421.

Jenkins, Katharine. 2018. Toward an Account of Gender Identity. *Ergo* 5.

Koch, Steffen. Forthcoming. The Externalist Challenge to Conceptual Engineering. *Synthese.*

Mallon, Ron. 2016. *The Construction of Human Kinds.* Oxford: Oxford University Press.

Manne, Kate. 2018. *Down Girl: The Logic of Misogyny*. Oxford: Oxford University Press.

Marques, Teresa. forthcoming. Amelioration vs. Perversion in *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability*, Teresa Marques and Asa Wikforss (eds.). Oxford: Oxford University Press.

Nado, Jennifer. 2019. Conceptual Engineering, Truth, and Efficacy. *Synthese* DOI: https:/doi.org/10.1007/s11229-019-02096-x

Podosky, Paul-Mikhail Catapang Podosky. 2018. Ideology and Normativity: Constraints on Conceptual Engineering. *Inquiry* https://doi.org/10.1080/0020174x.2018.1562374

Pollock, Joey. 2020. Conceptual Engineering and Semantic Deference. *Studia Philosophical Estonia* 12: 81 – 98.

Prinzing, Michael. 2017. The Revisionist Rubric: Conceptual Engineering and the Discontinuity Objection. *Inquiry*, DOI: 10.1080/0020174X.2017.1385522.

Sawyer, Sarah. 2020. Talk and Thought in *Conceptual Engineering and Conceptual Ethics*, Alexis Burgess, Herman Cappelen, and David Plunkett (eds.), pp. 379 – 395.

Sawyer, Sarah. Forthcoming. The Role of Concepts in Fixing Language. *Canadian Journal of Philosophy* 50 (5): 555 – 565.

Schroeter, Laura and Schroeter, Francois. 2020. Inscrutability and its Discontents. *Canadian Journal of Philosophy* 50 (5): 566 – 579.

Simion, Mona and Kelp, Christopher. 2020. Conceptual Innovation, Function First. *Nous* DOI: doi: 10.1111/nous.12302

Steinle, Fredrich. 2012. Goals and Fates of Concepts: The Case of Magnetic Poles in *Scientific Concepts and Investigative Practice*, Uljana Feest and Friedrich Steinle (eds.), pp. 105–125. Boston, MA: De Gruyter Mouton.

Strawson, P.F. 1963. Carnap’s View on Constructed Systems versus Natural Languages in Analytic Philosophy in *The Philosophy of Rudolf Carnap*, Paul Schilpp (ed.), pp. 502–518. La Salle: Open Court.

Thomasson, Amie. 2020. A Pragmatic Method for Normative Conceptual Work,’ in *Conceptual Engineering and Conceptual Ethics*, Alexis Burgess, Herman Cappelen, and David Plunkett (eds.), pp. 435 – 458. Oxford: Oxford University Press.

Valentini, Laura. 2012. Ideal vs. Non-Ideal Theory: A Conceptual Map. *Philosophy Compass* 7 (9): 654 – 664.

Vargas, Manuel. 2013. *Building Better Beings: A Theory of Moral Responsibility.* Oxford: Oxford University Press.

1. For instance, Barnes (2016, 2019) and Jenkins (2018). [↑](#footnote-ref-1)
2. Internalists might allow for more ‘meaning control’ (Koch forthcoming, p. 2). However, like Cappelen (2018), I think there is some uncertainty as to whether inner states are scrutable and willingly changed. See Pollock (2020) for a response to this. [↑](#footnote-ref-2)
3. Podosky (2018) is concerned with the means by which looping effects can be controlled; Fischer (2020) focuses on feasibility with respect to our ability to reason with engineers concepts; and Andow (forthcoming) argues that we should we the likelihood of success against the potential value of success. [↑](#footnote-ref-3)
4. Cappelen (2018) does state that no success conditions for conceptual engineering can be given in virtue of the terms stated in such conditions are subject to revision. [↑](#footnote-ref-4)
5. Cf. Haslanger on *epistemic amelioration* (2020, p. 242). [↑](#footnote-ref-5)
6. Other examples abound: Haslanger (2000), Thomasson (2020), Nado (2019). [↑](#footnote-ref-6)
7. In some cases, we might simply want to trash a linguistic practice in virtue of having a bad goal or purpose. [↑](#footnote-ref-7)
8. Laura and Francois Schroeter (2020) disagree with this. They argue that while we might not know all of the relevant facts that determine reference, this does not mean we are completely in the dark. And Steffen Koch says control is possible, but we need to recognize that it is a group-level project:

‘...even though reference change on causal theories of reference turns out to require a collective long-term effort, it is nevertheless something that we, as a linguistic community, can bring about willingly’ (forthcoming, p. 21) [↑](#footnote-ref-8)
9. Sawyer does give an account of ameliorative analyses of gender terms in which theorists are offering better theories about what the relevant subject matter is, *which includes trans women.* It is unclear whether this is the best construal. [↑](#footnote-ref-9)
10. Some like Fisher (2015), and perhaps even Prinzing (2018), tie function with meaning – the former offers a teleosemantics; the latter offers a view of concept identity. I follow Nado (2020) in thinking that meaning and function, at least in our effort to preserve continuity, should be treated separately. [↑](#footnote-ref-10)