Conceptual Innovation, Function First

Mona Simion

mona.simion@glasgow.ac.uk

Christoph Kelp

christoph.kelp@glasgow.ac.uk

Abstract

Can we engineer conceptual change? While a positive answer to this question would be exciting news for philosophy, there has been a growing number of pessimistic voices in the literature. This paper resists this trend. Its central aim is to argue not only that conceptual engineering is possible but also that it is not even distinctively hard. In order to achieve this, we will develop a novel approach to conceptual engineering, which has two key components. First, it proposes a reorientation of the conceptual engineering project away from fixing conceptual defects and towards bringing about conceptual innovation. Second, it offers a new account of when conceptual engineering is successful in terms of etiological functions. We then turn to the reasons that have motivated various forms of pessimism about conceptual engineering and show that, on our novel approach, none of them stands up to scrutiny.

1. Introduction

Conceptual amelioration talk is hip.[1] We want better ways of thinking about the world. Can we fix/replace our defective representational devices? Should we? Is this what philosophy is/should be all about? Extreme pessimists answer 'no' to some or all of these questions.[2] Cautious pessimists worry that concepts are hard to engineer because they are, to some extent or another,

[1] See e.g. (Author#2 2017, 2018), (Burgess, Cappelen and Plunkett forthcoming), (Cappelen 2018), (Clark and Chalmers 1998), (Greenough MS), (Haslanger 2000), (Leslie Forthcoming), (Plunkett and Sundell 2013), (Sharp 2013), (Thomasson forthcoming), (van Inwagen 2008). See (Cappelen 2018) for a great overview.
[2] See e.g. (Greenough MS).

out of our control.[3] Cautious optimists agree: it's hard. This shouldn't stop us, however: many things we do are hard, and this hasn't stopped us before.[4]

This paper defends extreme optimism about conceptual engineering. The optimist view has two key components. First, it is widely agreed in the literature that conceptual engineering is about repairing defective concepts. We think this unduly limits the scope of the endeavour. Accordingly, we propose a reorientation of the central focus of the project away from conceptual repair or and towards *conceptual innovation*. Of course, we can engineer new fixes for old defects. In this way, the old view is part of our proposal. Crucially, however, the new view goes well beyond this. In particular, it brings engineering new representational tools from scratch into the scope of conceptual engineering. The significance of this reorientation is hard to overestimate. As we will see later on, it contributes to providing attractive solutions to a number of the key problems that pessimists have been voicing about conceptual engineering.

The second key component for the optimist view is a specific account of when conceptual engineering is successful. When we engineer some product, we want it to do something for us. We carve out a function for our product that we want it to fulfil. Very roughly, our proposal has it that an engineering project is successful if the product comes into/remains in use because it fulfils the function[5] it was designed to fulfil reliably enough. The success condition for conceptual engineering is then straightforward. As we will argue in due course, this particular view of when conceptual engineering succeeds also helps answering worries of pessimists about conceptual engineering.

In conjunction, these two components serve to address the major obstacles that have been thought to stand in the way of us taking genuine charge in our efforts to bring about better concepts. Or so we will argue in Sections 4 to 6. Section 3 will develop the optimistic view of

[3] See e.g. (Chalmers 2011), (Eklund 2015), (Burgess and Plunkett 2013a and 2013b)
[4] See e.g. (Cappelen 2018).
[5] See also e.g. Haslanger (2000), (Brigandt 2010), (Thomasson forthcoming), and Richard (forthcoming) for function-friendly views. One extremely sharp question we have been asked in the past about our view (thanks a lot XXX for pressing us on this) is: how ambitious is the function-first claim? Is all normativity supposed to follow from function fulfillment? Given that the view is built on an etiological account of functions, where acquiring the function is itself dependent on first generating some variety of benefit, isn't this ultimately a good-first rather than a function-first view? We are sympathetic to the picture this worry paints: T-normativity is generated by T-functions, which in turn are acquired via T-value generation. That being said, the view we defend here does not rest on a particular account of functions, so neither does it rest on this value-theoretic picture. See fn. 8.

conceptual engineering in more detail. First things first, since the issue keeps on cropping up in conversation, we will say a more few things about what we take concepts to be (Section 2).

2. Concepts

An Englishman, his young Spanish niece and their dog are in the desert. After a long dry spell, they stumble upon an oasis. Having laid eye upon the refreshing liquid, all three immediately have the same thought: THAT'S WATER. Of course, the Englishman and his niece would give different linguistic expression to this thought. The Englishman would say "That's water" whereas his Spanish niece would say "Es agua". And the dog, not being in possession of a language such as English or Spanish, would not say anything at all. While the Englishman will also have the thought THAT'S H_2O, neither his niece nor their dog will do so. They don't have the cognitive sophistication to entertain this kind of thought, or in the case of the niece, at least not yet.

What are concepts? While we don't have a detailed account of concepts, we want to say the following by way of an answer: *concepts are the constituents of thought (contents)*. The above story helps us to home in on concepts by allowing us to see them in action. The thoughts THAT'S WATER and THAT'S H_2O are different thoughts. They must be. Otherwise, it couldn't be that the Englishman has both thoughts while the daughter and the dog only have one. How do they differ? The answer we suggest is that they have different constituents. In particular, the thought THAT'S WATER differs from the thought THAT'S H_2O in that the former features the concept WATER as a constituent, whereas the latter features the concept H_2O. The reason why all three can have the thought about water but only the Englishman can in addition have the thought about H_2O is that while all three possess the concept WATER, only the Englishman possesses the concept of H_2O.

Concepts are also different from words and other forms of linguistic expressions. Again, they must be. Otherwise, it couldn't be that the Englishman and his Spanish daughter have the same thought. Moreover, concept possession does not require the possession of a language such as English or Spanish. Otherwise, it couldn't be that the dog and his masters have the same thought. At the same time, concepts share some important properties with linguistic expressions,

to wit, they have intensions and extensions. We take both extensions and intensions to be relative to circumstances of evaluation, e.g. worlds or world-time pairs. The extension of a concept at a circumstance is the set of all and only things to which the concept applies at that circumstance of evaluation (CE). For instance, the extension of the concept COW at (W, t) is the set of all and only cows at (W, t). The intension of a concept is a function from CEs to extensions of the concept at the input CE.[6]

3. Conceptual Engineering: the Optimistic View

3.1 The Innovation Component

The literature on conceptual engineering has been focusing largely if not exclusively on conceptual repair. By way of evidence, consider the following passage from Greenough:

> Conceptual defects are many and varied. Concepts (and terms) can be incomplete ('open-textured'), confused, unsatisfiable, vague, or inconsistent. They can be too inclusive, too narrow, or simply empty. They can be too complex, too simple, or not fit to feature in any useful explanation; they can be superseded, tired, hackneyed, or systematically misapplied. They can be too parochial, too elitist, or too recondite. They can be loaded with inappropriate connotations, bad ideological baggage, or serve as ongoing devices for deceit, discrimination, or oppression. A concept may be flawed on more than one dimension—broken in many different ways. Conceptual Engineering, as a result, is a multifarious business (Greenough MS, 3).

[6] This rough characterisation is compatible with all the leading views about the structure of concepts. Concerning structure, our characterisation is compatible with the view that concepts have definitional structure, prototype structure, theory structure and atomic structure. We do not mean to take a stance on either of these issues here. That said, it is tempting to think that concepts are constituted by principles governing their application. After all, this appears to be something that all of the rival views on the structure of concepts can agree on. Where they disagree is on how these principles must be related to one another in order to constitute a concept. Moreover, our view is compatible with a range of attractive views about concept possession. For instance, the view is compatible with the view that to possess a concept is to be in a kind of mental state, e.g. the state of knowing the principles that constitute it, or that it is an ability, e.g. the ability to correctly apply the concept by means of its constitutive principles (perhaps in suitably favourable conditions). That said, not much hangs on a particular view of putting meat on the bones of the above rough characterisation of concepts or the above views of concept possession, at least as far as our purposes are concerned. Rather, our point here is that what little we do say about concepts can be developed into a detailed account of concepts and their possession in a number of promising ways. And that should make it all the more acceptable.

In a similar vein, Herman Cappelen's has recently proposed a taxonomy of conceptual engineering which has, at its very core, a variety of different types of conceptual deficiencies and a number of ways in which we may venture to fix them. What's particularly telling about Cappelen's taxonomy is that it is reverse engineered from extant proposals for conceptual engineering in the literature. All of this serves to provide support to the idea that there is a general consensus among those working on conceptual engineering, to wit, that *it is about fixing defective concepts*.

At the same time, there is reason that this way of thinking of conceptual engineering unduly limits its scope. To see this, note that it is widely agreed in the theory of normativity that in order to justifiably embark on a certain project, such as a conceptual engineering project, all that's needed is improvement, not fixing a defect (Author#2 2017). Or to be more precise, for all phi, what T-justifies phi-ing is a T-type improvement, not fixing a T-type defect. Of course, fixing a defect of a certain sort is one way of bringing about a corresponding type of improvement. But it is not the only way. If doing something will improve things in a certain way, we'll have reason to do that thing, even if we do not thereby fix a defect. For instance, if it's morally better for me to give money to charity than to not give money to charity, than I am justified in giving money to charity. It need not be that I am thereby remedying some moral defect. If it is prudentially better that I go to Mary's party tonight, than I am prudentially justified in going to Mary's party tonight. And so on.

If that is the case, however, it is just not clear why those interested in conceptual engineering should restrict their focus on fixing defective concepts, rather than on improving the world of concepts. For instance, say that there was nothing wrong with our concept WOMAN, semantically, morally, politically or otherwise; say that it is a perfectly coherent concept, and its current shape has zero detrimental effects on women's moral, political or epistemic life. Say, however, that it could be engineered such as to substantially improve women's lives. Would it not be worthwhile to attempt to do so? We take it to be pretty clear that the answer here can only be 'yes'. But for those who remain in doubt, it may be worth considering briefly a closely related branch of engineering, i.e. social engineering. In particular, it may be worth asking the question as to whether we need our social institutions to be somehow deficient in order to be

justified in proposing engineering projects, or whether it is it enough if changing our social institutions will have a positive effect on our lives? There can be no doubt as to what the right answer must be. And, of course, there is every reason to think that the same goes, mutatis mutandis, for conceptual engineering.[7]

The upshot of this is a fairly optimistic picture concerning the aim of conceptual engineering: we should broaden up! Ambitions of conceptual engineering need not be motivated by defects in our representational devices; proposals for improvements in the world of concepts will do just as well. Accordingly, what we would like to propose is a reorientation in focus of the conceptual engineering project: what's central is *conceptual innovation* not conceptual repair. It may be worth noting that this reorientation does not exclude extant proposals for conceptual engineering. After all, they all involve proposals for conceptual innovation, even if innovation consists in the modification of an existing concept and is sold as a form of conceptual repair. This is the first constructive component of our optimistic view of conceptual engineering.

Why does this matter? To preview our answer, the reason is that it helps defuse some important objections that have been levelled against the very possibility of conceptual engineering. While we will get back to this in due course, we'd like to ask you to bear with us for the time being. What we want to do first is develop the second constructive component of the optimistic view, which is to develop an account of when a conceptual engineering project is successful.

3.2 The Function Component

Consider the heart. It is a paradigm case of a functional trait. Its function is to pump blood. What are functions? According to the leading account in the philosophy of science, the etiological theory of functions,[8] functions turn on histories that explain why the item exists or operates the way it does. In the case of the heart, tokens of the type pumped blood in our

[7] Of course, strictly speaking, T-justification does not supervene on T-obligation, but rather on T-permissibility. The question, however, remains: why is the conceptual engineer is in the business of ameliorating concepts that are in *need* of amelioration, rather than ameliorating concepts that *can* be ameliorated.

[8] Defended by people like David J. Buller (1998), Ruth Millikan (1984), Karen Neander (1991), Peter Godfrey-Smith (1994) and, last but not least, Larry Wright (1973). The etiological theory of functions is, by far, the most widely endorsed view in the literature, due to its normative import. That being said, the view defended here does not rest on rest on the etiological view: any account of functions that delivers the highly desirable result that functions generate norms (i.e., that there is such a thing as a malfunctioning/properly functioning heart) will do for our purposes.

ancestors. This was beneficial for our ancestors' survival, which explains why the heart has continued to exist in creatures like us. As a result, the heart acquired the etiological function (henceforth also e-function)[9] of pumping blood.

The heart acquires its e-function via natural selection over generations. Not all functional traits follow the model of the heart: there will be cases where a requirement of selection over generations for function acquisition will seem implausibly strong (Sosa 1993). The paradigmatic case is that of beneficial macro-mutations, so-called 'hopeful monsters' (Graham 2014, 30). Most mutations are harmful (think of extreme birth defects); once in a while, though, a happy accident happens: someone is born with an almost entirely new trait or organ, very different in kind from its ancestral trait, which actually benefits the recipient. Since they are mutations, they don't have an evolutionary history; they are 'first generation' traits. Still, they can have functions. What matters is that the existence/continuous existence of a trait is explained via a history of positive feedback:

> [Etiological] functions arise from consequence etiologies, etiologies that explain why something exists or continues to exist in terms of its consequences, because of a feedback mechanism that takes consequences as input and causes or sustains the item as output (Graham 2014, 35).

Some things have *designed functions* (d-functions). The dishwasher is a paradigm example of an item with a d-function. Its d-function is to clean dishes. Items with d-functions have their d-functions in virtue of the intentions of the designer. The reason why the dishwasher has the d-function of cleaning dishes is because the inventor of the dishwasher intended it to clean dishes. Crucially, e-functions are different from d-functions. For starters, there are things that have e-functions but not d-functions. The heart is a clear example here. At the same time, there are things that have d-functions but not e-functions. The reason for this is that e-functions require a history of success. The heart could only acquire the e-function of pumping blood because token hearts successfully pumped blood in the past. Exhibits in the Museum of Failure have d-

[9] For applications of the etiological account of functions to epistemology and language, see e.g. (Author#1 2018), (Author#2 2017, 2018,), (Author#2 & Author#1 2016), (Graham 2012), (Millikan 1984).

functions but no e-functions: they just didn't work.

At the same time, many items with d-functions also acquire e-functions. Consider, in particular, new products which are launched on a competitive market. These products have d-functions. They are meant to do something. If they are successful, they will in addition acquire a certain kind of e-function. If all goes well, these products will initially be bought to do what they are designed to do and subsequently will continue to be in demand (if all goes well) in virtue of the fact that they did the things they were designed to do and that this was beneficial to consumers. But given that they continue to be in demand, they will continue to be produced and bought. It is now easy to see that we have exactly the kind of feedback loop going that is characteristic of e-functions. In this way, d-functional items may acquire e-functions. In fact, a plausible aim of designers who develop new products to be launched on a competitive market is for the d-functions of their products to turn into e-functions of the sort just mentioned. By the same token, one important success condition for a d-functional item is for its d-function to be converted into the relevant kind of e-function.

On the present view, conceptual engineers are in the first instance designers. They either develop new d-functions for existing concepts or else new concepts with new d-functions altogether. In the case of concepts, the d-function will consist in thinking about the world in a certain way. Moreover, these concepts are then launched on a competitive market of concepts. If all goes well, they will be used to do what they are designed to do and they will continue to be used in virtue of the fact that the way of thinking about the world they made available was beneficial to users. Again, we have exactly the kind of feedback loop going that is characteristic of e-functions. It will come as no surprise, then, that we want to suggest that a plausible aim of conceptual engineers is for the d-functions they develop to turn into the kind of e-function just described.

This completes the constructive part of our defence of the optimist view about conceptual engineering. To repeat, there are two central components: the first is a reorientation in focus from conceptual repair to conceptual innovation and the second is a view of what it takes to succeed in conceptual engineering: for the proposed d-function to convert into a certain kind of e-function.

4. All Out Pessimism

In the next two sections we will look into a battery of arguments that pessimists about conceptual engineering have adduced for thinking that it is either not possible at all or else that it is distinctively hard. We will show how the optimistic view developed above can avoid the problems on all sides.

4.1 The Strawsonian Worry

The issue of whether (a philosophically useful) form of conceptual engineering is even possible has a venerable history in analytic philosophy, dating back at least as far as Carnap and Strawson. Carnap was one of the first to explicitly countenance a form of conceptual engineering, which he called explication. Here is Carnap:

> The task of making more exact a vague or not quite exact concept used in everyday life or in an earlier stage of scientific or logical development, or rather of replacing it by a newly constructed, more exact concept, belongs among the most important tasks of logical analysis and logical construction. We call this the task of explicating, or of giving an *explication* for, the earlier concept … (Carnap 1947, 7-8)

Carnap also clearly thought that this kind of project was possible. In this way, he was an early optimist about conceptual engineering.

Strawson on the other hand is more cautious. He worries that the form of conceptual engineering Carnap had in mind may ultimately be tantamount to simply turning our backs on the philosophical problems we set out to solve in the first place:

> [T]ypical philosophical problems about the concepts used in non-scientific discourse cannot be solved by laying down the rules of use of exact and fruitful concepts in science. To do this last is not to solve the typical philosophical problem, but to change the subject. (Strawson 1963, 506)

While Strawson worries about aiming to solve philosophical problems by engineering scientific concepts, it is easy to see that the problem generalises to doing so by engineering philosophical concepts. Changing the concepts still seems tantamount to simply turning one's back on the old philosophical problems. After all, the old problems are stated in terms of the old concepts.

We have two suggestions by way of response. First, it may be worth noting that walking away from a problem isn't always a bad thing. On the contrary, sometimes it is exactly what needs to be done in order to make progress. For instance, Ptolemy might have charged Copernicus with simply turning his back on the problem he was interested in, to wit, explaining planetary motions within a geocentric worldview. Even if this may well be true, there can be no question that what Copernicus did was exactly what needed to be done to make progress in astronomy. Likewise, we want to suggest, sometimes it may be that turning our backs on problems stated in terms of certain concepts may be exactly what needs to be done in order to make progress in philosophy.

Second, the optimistic view we have outlined allows for conceptual engineering to be done without abandoning old problems. The key reason for this is the proposed reorientation from the repair view to the innovation view of conceptual engineering. Of course, according to the repair view, conceptual engineering is all about getting rid of defective concepts. By the same token, it commits us to abandoning philosophical problems stated in those concepts. However, once we move from repair to innovation, it is clear that conceptual innovation can happen without getting rid of the old problematic concepts. By the same token, this approach does not require us to abandoning philosophical problems stated in terms of them. On the contrary, those interested in them can continue to work on them (using the old concepts) if they are so inclined.

4.2 The Proliferation Worry

One way of thinking about Strawson's worry is that conceptual engineering leaves us philosophers with too little work to be done. The proliferation worry in a way presses in exactly the opposite direction. The thought here is that conceptual engineering puts too much work on our plates. To see why, note that concepts are connected to one another. If you revise a concept

or replace it by another one, it may well be that you need to work on other concepts to which it is related by conceptual principles. Consider, for instance, the concept of truth. Here is Greenough on what would have to happen if we wanted to replace the concept of truth by a different concept:

> [N]ot only is the concept of truth constitutively linked to the concepts of knowledge, provability, assertion, and belief, it is also inextricably linked to myriad other (philosophical) concepts such as the concepts of inquiry, objectivity, reality, world, judgment, evidence, justification, confirmation, probability, fact, being, truth-value, truth-bearer, concept, analyticity, reference, denotation, satisfaction, truth-condition, intension, extension, meaning, content, proposition, representation, necessity, possibility, contingency, externalism, reliability, and more. In turn these concepts are constitutively linked to a wider class of concepts which may well encompass, in the end, all concepts of central philosophical interest. *All* these concepts must be replaced too, together with words we use to pick them out. (Greenough 2017 Chapter 3, Section 5.3)

What Greenough's example of engineering the concept of truth forcefully indicates is that replacing the concept of truth may set in motion an entire avalanche of needs for conceptual engineering which, besides questions of feasibility, raise clear questions of desirability.

Our response is once more twofold. First, it is not clear that the proliferation worry will affect all forms of conceptual engineering equally, even those that aim at repair. True, those forms of conceptual engineering that are motivated by escaping paradox will be fairly likely to face the proliferation problem, especially if the concept to be engineered is a central philosophical concept such as the concept of truth. However, there are forms of conceptual engineering that seem much less susceptible to this kind of worry. Consider for instance Author #2's (2017) proposal to engineer the concept of epistemic norm. The defect that Author #2 diagnoses is that the concept of epistemic norm does not do well on prior plausibility, i.e. it doesn't fit with how we individuate norms in the general theory of normativity. Here the defect is that the concept doesn't hook up with other concepts in the way it ought to and the solution is

to engineer it in such a way as to do so. It is far from clear that this form of conceptual engineering faces the proliferation worry.

Second, the optimistic view can circumvent this worry in much the same way as it could avoid of the Strawsonian worry. Once we are clear that conceptual engineering has conceptual innovation at its heart, there is little reason to think that the proliferation worry tears the entire project asunder. After all, conceptual engineering can proceed by producing something new, whilst leaving the old concepts in place.

4.3 The Inductive Worry

The core idea of the inductive worry is that we have good inductive reason to think that conceptual engineering can't work. The thought here is that there exists no successful case of conceptual engineering in the history of mankind. What's worse, this is not for want of trying. Totalitarian regimes have time and again embarked at least on projects of linguistic engineering. However, they have all eventually failed. By the same token, we have reason to think that conceptual engineering is hopelessly utopian. Or so the thought goes.

As a first observation, note that, at best, the inductive worry applies to the repair view of conceptual engineering. Once you broaden your view of conceptual engineering to include all forms of conceptual innovation, a whole host of success stories comes into view. The history of science is filled with successful instances of conceptual engineering. Concepts such as electron, hydrogen or DNA have not been with us forever. Rather they were engineered by scientists in their attempt to make sense of the world around us.

That said, we are not all that convinced that the only relevant evidence comes from failed projects of linguistic engineering by totalitarian regimes. On the contrary, we think that there are also successful cases of linguistic engineering. Take, for instance, job titles. Especially in recent years old terms have more and more frequently been replaced by newer ones. For instance, what used to be 'marketers' are now 'account managers', 'hackers' are 'developers' and 'shop clerks' are 'sales representatives'. It may also be worth noting that the optimistic view of conceptual engineering can explain the differences. In the case of totalitarian regimes, the value that use of a new term generated and that kept it alive consisted solely in avoidance of punishment by the

regime. Once the threat of punishment is not in the picture, the incentive to use the term diminishes greatly or disappears altogether. No surprise, then, that in contexts in which there is little reason to fear the consequences or once the regime has broken down, the terms will no longer be used. The case of the job titles is importantly different in this respect. The value the use of these terms produce include that employers and co-workers signal recognition and respect for the bearers of these titles that is they feel better about themselves and are more motivated at work. We submit that this is exactly the kind of value that will not lead to the kind of failure in the long run that has riddled totalitarian regimes. (Of course, we may engineer even better job titles which go on to replace the current ones, but that's an indication of successful engineering, not of ultimate failure.) By the same token, the inductive evidence is not as clear cut as those who have pressed the inductive worry would have us think. On the contrary, if anything, a careful look once again gives us reason for optimism about conceptual engineering.

5. Cautious Pessimism

At this stage, even some the pessimists might be inclined to grant us that a certain form of conceptual engineering is possible, to wit, one in which we carve out a concept that allows us to think about the world in a new way and perhaps coin a new term for it. If this way of thinking about the world is good for us and the new term sufficiently catchy, the concept may acquire the right form of e-function. The conceptual engineering project will have been successful. But even if this form of conceptual engineering is successful, pessimists might still hold on to the idea that there is an important other form of conceptual engineering that remains problematic. This is the form of conceptual engineering in which we try to engineer *change in an existing concept* (henceforth 'conceptual change engineering'). What's more, this is exactly the form of conceptual engineering that philosophers have hitherto pursued. So even if conceptual engineering is possible, the pessimistic might still win the day in the sense that extant incarnations of conceptual engineering are bound to fail.

The central driving force for this worry is that conceptual change engineering is hard, if not impossible, to combine with semantic externalism. According to semantic externalism (e.g. Burge 1979), for instance, features of the external environment we find ourselves in partly

determine the semantic value of our concepts. The relevant elements of the external environment include experts in the community, the history of use going back to the introduction of a term, complex patterns of use over time, and what the world happens to be like (independently of what we believe the world to be like). In a nutshell, the externalist worry is this: Given that semantic externalism is true, it is hard to see how we could have the kind of control over our concepts that engineering would appear to require us to have. Here are Burgess and Plunkett (2013) voicing this worry:

> The textbook externalist thinks that our social and natural environments serve as heavy anchors, so to speak, for the interpretation of our individual thought and talk. The internalist, by contrast, grants us a greater degree of conceptual autonomy. One salient upshot of this disagreement is that effecting conceptual change looks comparatively easy from an internalist perspective. We can revise, eliminate, or replace our concepts without worrying about what the experts are up to, or what happens to be coming out of our taps. From the externalist's point of view, however, conceptual revolution takes a village, or a long trip to Twin Earth.[10] (Burgess and Plunkett 2013, 1096)

5.1 Engineering Biological Change

No one would want to deny that we have less control over semantic facts than, say, over whether to raise our arms, especially if semantic externalism is true. That said, as we will argue momentarily, we don't need this kind of control over semantic facts for conceptual change engineering to be possible. In fact, what we have is all we need. To see this, we will first consider

[10] As the passage already suggest, Burgess and Plunkett, they think the lesson to be learned for the conceptual engineer is to embrace semantic internalism. Unfortunately, many, ourselves included, think that embracing semantic internalism simply won't do the trick for conceptual change engineers as considerations like the following forcefully indicate:

> [M]ost or even all speakers of the language can believe that a predicate F applies to an object, o, but be wrong. They can all want o to be in the extension of F, but wanting o to be F doesn't make it so. They can all be disposed to apply F to o even though o isn't F. Humpty Dumpty was wrong: believing and wanting words to mean something doesn't make it so (Cappelen 2018, 65).

Cappelen point here is about words. Crucially, however, it is quite plausible that the lesson generalises to concepts. After all, thought is often enough linguistic in shape. When it is, the fact that the semantic facts about words do not supervene on our internal states only implies that, at the very least, which concepts we are using in linguistic thought does not supervene on our internal states only. And if that's right, then semantic internalism about concepts simply does not seem to be a viable option for conceptual engineers.

a case for which it is plausible that we have no more control over relevant facts than we do in the semantic case.

Recall that biological traits can have e-functions. For instance, the heart has the e-function of pumping blood. E-functions can change. One way in which this happens is when the relevant habitat changes and, as a result, the trait no longer fulfils its function reliably enough when functioning normally. Here the trait comes under pressure to change its modus operandi, i.e. the way it functions. And if all goes well the new modus operandi will reliably fulfil the relevant function in this new habitat. The trait has adapted to the new environment. When it fails to do so, the trait gradually loses its function and becomes a vestigial trait or else becomes extinct altogether.

Here is another way in which a trait's e-function may change. The trait is in a fairly stable habitat, where very minimal change in the relevant conditions affects the trait's reliability in function fulfilment little to not at all. Even so, the trait may start functioning abnormally in normal conditions and, crucially, this is more beneficial to the organism, i.e. the abnormal way leads to an increase in how reliably the trait fulfils its function. With time (feedback) the formerly abnormal way functioning becomes the new normal way of functioning. Once again adaptive change has occurred.

Can we control changes in biological e-functions? On the face of it, the prospects for this look rather bleak. In particular, what we'd like to highlight is that they look about as bleak as in the case of concepts if semantic externalism is true. In both cases, it would seem we do not have the kind of control that would be needed to achieve this.

What we want to argue is that there is reason for optimism after all. We can engineer adaptive changes in both biological traits and in concepts. In what follows we will first take a closer look at the biology case. We will show that we can bring about adaptive changes and how this can happen. Moreover, we will argue that if we can bring about these changes, we can also engineer them. Finally, we will argue that the same goes, mutatis mutandis, for the case of concepts.

Can we bring about adaptive change in biology and if so how? We think so. Here's the key idea, in a nutshell. We are part of the relevant habitats. That's why we can act on these

habitats and bring about changes in them. As a result, we can bring about exactly the kinds of changes in these habitats that generate pressure to adapt. In this way, the possibility of engineering adaptive change comes alive.

By way of more concrete evidence, consider the case of the peppered moth. Light-coloured pepper moths were the norm before industrial evolution. In fact, the dark-coloured (or melanic) variety of the peppered moth was not even known before 1811. During the industrial evolution, after a field collection in Manchester, the frequency was found to have increased dramatically. By the end of the 19th century, the dark moth had taken over almost entirely, with a record of 98% in 1895. Bernard Kettlewell was the first to investigate the evolutionary mechanism behind peppered moth adaptation, between 1953 and 1956. He found that a light-coloured body was an effective camouflage in a clean environment, such as in Dorset, while the dark colour was beneficial in a polluted environment like in Birmingham. This selective survival was thought to be due to birds which easily caught dark moths on clean trees, and white moths on trees darkened with soot.

The peppered moth is not an isolated example of biological change where our intervention on the environment carries explanatory salience. Bedbugs are another example: since we have used an abundance of chemicals to get rid of them, bedbugs have developed thicker shells and tougher nerve endings.[11] Climate change, insofar as we agree that we played our role in this, affected both behaviour and traits in an overwhelming number of species.[12] In all of these cases our input into the adaptive process is explained by our influence on the habitat. We change the habitat, the relevant trait needs to adapt to the new one. Even though we cannot directly change biological traits as we please, we can affect them so by acting on their habitat, and thereby function fulfilment, thereby triggering the corresponding adaptive change.

Of course, while these cases clearly indicate that we do bring about biological change, bringing about change is one thing, engineering it is another. In particular, it might be thought that to engineer adaptive change, not only must it be possible for us to generate changes in habitat, we must also be able to control changes in the right way. Even if we agree that we can bring about changes in natural habitats, this is not to say that we can control these changes in a

[11] See e.g. http://u.osu.edu/bedbugs/
[12] See e.g. (Monroe et al. 2018).

way that would be required for engineering these changes. So, we haven't yet shown that we can engineer biological change.

On the upside, once we recognise that we can bring about biological change, it should not come as a surprise that we can also engineer it. In fact, once again, there is excellent reason to think that engineering biological changes is a reality. Consider poliovirus, which, not too long ago, was prevalent across the globe. Thanks to human ingenuity we designed a vaccine and given it to children across the world with the result that it has now been eradicated in nearly all of its former habitats. We engineered biological change through strategic interventions in the relevant habitat.

But what about biological change that is initiated not by a change in the habitat but by a change in the modus operandi of the trait? Can this kind of change be engineered also? While for the longest time it would have been thought that the answer to this question has to be no, at present there is excellent reason to say that it is yes. In fact, there is reason to think that this kind of engineering biological change is also a reality. Genetic engineering directly manipulates organisms' genes by the use of biotechnology. Although quite a complicated affair, this technique has become quite mainstream. When the changes we engineer become adaptations, i.e., when they do well/better than the original trait in their habitat, they acquisition etiological functions and get reproduced over generations.[13]

5.3 Engineering Conceptual Change

It comes to light that, despite the fact that biological traits depend heavily on the world and despite the fact that we don't have the kind of control over biological traits that we have, say, over whether to raise our arms, both human induced biological changes and engineered biological changes are not only possible but also actual. The last step in our response to the cautious pessimist worry is that engineering conceptual change is possible also for much the

[13] It may also be worth noting that engineering biological changes isn't uniquely difficult, in comparison with engineering changes in artefacts or social structure. To succeed in the latter will often be no less difficult. Sure, change can be effected overnight. We can vote to change our political structures, and the owner of some car company may just decide to only produce electric cars next year. However, it's not clear that, without these changes scoring well in the relevant environment, they will persist. Or, to put it in functional terms, insofar as the new social structure/car does not deliver the relevant benefit, and does not, thereby, fulfil its function well/reliably, the change is likely to not be an adaptive one, and thus be abandoned. Classical examples of such failed attempts at change in social structures are Eastern European communist regimes. Classical examples of failed attempts at changes in artefacts are displayed in museums all over the world, with the 'Museum of Failure' in Sweden leading the charts.

same reasons as it is possible to engineer biological change. It is possible for us to bring about conceptual change. And if so, there is every reason to think that it is also possible for us to engineer conceptual change.

In a recent paper, Cian Dorr and John Hawthorne discuss the example of SALAD (Dorr and Hawthorne 2014). Not too long ago, SALAD picked out mixture of leaves of sorts with a touch of dressing. Dorr and Hawthorne note that we now find it unproblematic to apply SALAD to "various warm leaf-free concoctions" (2014: 284). In fact, today, there is little to no reason to think that a full chicken on top of three slices of cheese could not fall under the concept SALAD.[14]

Conceptual change is adaptive change. That is to say, when conceptual change happens what changes is the e-function of the concept. To see the plausibility of this, consider the following plausible story about how SALAD changed: we got richer and we started being more concerned with our health and looks. Salads became more and more popular. Food providers – be they grandma or the local fancy restaurants – rushed to accommodate the demand. First came variation: increased demand for a particular type of good has this effect: it generates increased offer from more providers, who then compete for the market by variating on the properties of the relevant commodity in an attempt to create the most successful product. Second, a particular type of variation was encouraged: variation towards a more nutritious product. After all, in times where it is fashionable to stick to a mere salad for lunch, it had better be a wholesome dish. A change in the habitat of the concept SALAD led to a change in the concept of SALAD.

Finally, let's return to the question whether we can also engineer conceptual change. Recall the biological case, which is a realm of phenomena that is very strongly dependent on the world and over which, on the face of it at least, we may seem to have very little control, if any at all. Even so, we not only can bring about change but we can also engineer it. In fact, there are concrete examples both of actual humanly induced biological change and of engineered biological change. We have also seen how we can and do achieve this, to wit, by acting on the habitat, strategically in the case of bio-engineering, with the result that pressure towards adaptive

[14] Note also that SALAD isn't very distinctive in these respects. Changes in concepts, to some extent or another, happen all the time: consider, e.g., BOOK, WATCH, LUNCH, HEALTHY, MARRIAGE, PERSON and RAPE, to name but a few (Cappelen 2018, 32).

change is generated. What the above considerations show is that, in the conceptual case, we can bring about conceptual change in much the same way, i.e. by acting on the environment, with the result that pressure towards adaptive change is generated.

Given that this is so, is there still any good reason to think that it should not also be possible to engineer conceptual change by intervening strategically in the habitat of a concept? We take it to be beyond question that the answer to this question is no. In particular, we take it to be clear that the fact that, on the face of it, we have very little to no control over semantic facts provides no good reason for thinking that we cannot engineer concepts in this way, just as the fact that, on the face of it, we have little to no control over biological facts provides no good reason for thinking that we cannot engineer biological traits.

Finally, what about the second form in which adaptive change can take place, i.e. through a change in the modus operandi of the relevant concept? Again, there is every reason for thinking that this can be done. Those engaging in the practice of conceptual engineering can and have proposed changes in the modus operandi of various concepts, including by advancing modified definitions of our concepts (e.g. Haslanger on the definition of the concept of woman), by proposing changes to what can be in their extensions (e.g. homosexual being able to be in the extension of the concept of marriage) and so on. Just as in the biological case, when the changes we engineer become adaptations, i.e., when they do well/better than the original concepts in their habitats, they acquisition the relevant e-functions in which case the engineering project will have been successful.

6. Anti-Optimism

Let's say that the last two sections did successfully defuse the central worries that pessimists have harboured about conceptual engineering. If so, there is reason to believe that the optimistic view we developed in section 3 does carry promise. Before closing, we would like to look at a couple of worries that target the optimistic view in particular. After all, if our optimistic view turns out to not to be viable at the end of the day, our attempt to argue for optimism about conceptual engineering of any sort on its basis is bound to failure also.

6.1 Survival of the Fittest?

Especially in view of the functionalist component of our view, it might be thought that what we are advocating here is a form of conceptual engineering according to which practitioners generate proposals regarding conceptual innovation and wait for the fittest concept to survive. However, it might be thought that there are a number of problems with this view. First, it's just not clear that what survives always has overall beneficial effects. For instance, the python may survive in the everglades and destroy the rest of the environment and creatures. Yet we wouldn't want to say that this has beneficial effects overall. Second, it's not clear that the best concept will indeed win out. For instance, most experts would agree that Betamax was the better product than VHS but the latter emerged as preeminent. Third, it's just not clear that survival of the fittest is the right model for all ways in which concepts may get uptake, especially in cases in which concepts are imposed in a top-down manner (e.g. psychiatrists writing the DSM).[15]

The key to our response to all of these worries is that the e-functionalist component of the optimistic view is a view about one way for a conceptual engineering to be successful, not a view about what must be the case for conceptual engineering to be successful or about how engineered conceptual change must take place. Conceptual engineers whose engineered concepts acquire the right sort of e-function will have succeeded in their projects, just as designers of more familiar sorts will have succeeded if their products acquire the right e-function. Just as with more familiar forms of design projects, it is entirely compatible with this (i) that the successfully designed product/engineered concept is detrimental overall, (ii) that there are better products/concepts out there and (iii) that products/concepts may come to be used in different ways.

One thing that the above considerations do highlight is that not every successful instance of conceptual engineering will be an instance of conceptual amelioration in the sense that it will make the world an overall better place. Now think that this is entirely unsurprising, since the same is true of more familiar forms of design. Crucially, even though successful product design, say, need not amount to an overall improvement of the world, it can do so and it often enough does. And there is no reason to think that the same should not hold, mutatis mutandis, for

[15] Thanks to YYY for pressing us on these issues.

conceptual engineering. At the same time, it is important to keep in mind that the main worries about the viability of conceptual engineering all effectively target the possibility of conceptual engineering. Our aim was defend the viability of conceptual engineering by showing that conceptual engineering can be done and successfully so.

Of course, conceptual engineers are, typically at least, not only interested in placing a new concept on the concept market. Rather, they will also expressly want to improve the world in some way and they also want to avoid making it a worse place. Another important point the above considerations highlight is that conceptual engineering had better not take place in a void, especially if it is to be part of a broader aim to not only engineer new concepts and get people to use them, but to thereby make the world a better place. Most pertinently, there is an important role to be played in the broader endeavour for what has become known as conceptual ethics, which is tasked with looking after the normative wellbeing of our concepts, as it were.[16]

6.2 Can there be a Success Condition for Conceptual Engineering?

We offer a success condition for conceptual engineering. But is this even possible? Herman Cappelen thinks the answer to this question is no. Here goes:

> If you expect a theory that provides necessary and sufficient conditions for when successful conceptual engineering has happened (or an instruction manual for how to do it), this book will disappoint. *I argue that no such theory can be given*: even the success condition for conceptual engineering are up for grabs – revisionism is happening also at the meta-level – the rules for conceptual engineering are constantly engineered.

If you are serious about conceptual engineering, the very concepts at issue in the success condition are up for grabs. As a result, the project of giving success conditions for conceptual engineering is hopeless. Or so Cappelen seems to think.

We think that there are at least two problems with this argument. First, even if it is true that all concepts are up for grabs in the sense that we should inquire into whether they require

[16] See (Burgess and Plunkett 2013a, 2013b).

engineering, it does not follow that all concepts do require engineering. It might be that some concepts turn out just fine on inspection. The claim that all concepts do require engineering itself affords argument and, on the face of it at least, does not look altogether promising. In the absence of such an argument, it seems to be a fair methodological assumption that we are entitled to use extant concepts in theorising at least given that we have no reason for thinking that they require engineering. Since Cappelen has not provided reason for thinking that all concepts are in need of engineering, nor that any of the concepts at issue in our success condition for conceptual engineering are, we should be in the clear, at least for the time being.

Second, even if some of the concepts at issue in our success condition do require engineering it does not follow that our proposal does not continue to give a sufficient condition for success in conceptual engineering, after the relevant concepts have themselves been engineered, that is. After all, post engineering, we may still end up with a condition for success in conceptual engineering that is not only sufficient but also more inclusive in the sense that it rules in instances of conceptual engineering as successful that the old one wouldn't have without failing to rule in any instances that the old one would have. What's more, even if the results of the needed engineering are not that favourable for us, it may well be that our success condition still allows us to identify various things as successful instances of conceptual engineering. To see what we have in mind, consider: at some point in the past, the concept of fish included whales and mammals. Let's agree that, at that point, the concept of fish was in need of engineering. Does it follow that our ancestors who employed this concept of fish could not use this concept to identify various things as fish, including their pet goldfish, the seabreams they had for dinner, etc.? The answer here is clearly 'no'. And, of course, the same may very well hold, mutatis mutandis, for the concepts at issue in our success condition for conceptual engineering. If that's right, there is little cause for concern even if some of the concepts at issue in our success condition are in need of engineering, again at least for the time being.

7. Conclusion

In sum, there is reason for thinking that the prospects for conceptual engineering are bright. Once we embrace the proposed reorientation of the focus of conceptual engineering from

conceptual repair to conceptual innovation and the proposed success condition for this enterprise, it comes to light that conceptual engineering is very much a feasible project. Our take home message, then, is to no longer fret about making the world of concepts a better place; just go ahead and do it!

References

[Note: this is a fast-moving field, many of the works referenced in the paper are not yet published]

Author#1 (2018). Author's Work.

Author#2 (2018). Author's Work.

Author#2 (2017). Author's Work.

Author#2 & Author #1 (2016). Authors' work.

Brigandt, Ingo (2010). The epistemic goal of a concept: accounting for the rationality of semantic change and variation. *Synthese* 177 (1): 19– 40.

Buller, D.J. 1998. Etiological theories of function: A geographical survey. *Biology and Philosophy*, 13, 505–527.

Burge, Tyler (1979). Individualism and the mental. *Midwest Studies in Philosophy* 4 (1): 73– 122.

Burgess, Alexis, Cappelen, Herman, and Plunkett, David (eds.) (forthcoming). *Conceptual Ethics and Conceptual Engineering*. Oxford University Press.

Burgess, Alexis, and Plunkett, David (2013a). Conceptual ethics I. *Philosophy Compass* 8 (12): 1091– 101.

Burgess, Alexis, and Plunkett, David (2013b). Conceptual ethics II. *Philosophy Compass* 8 (12): 1102– 10.

Carnap, Rudolf (1947). *Meaning and Necessity*. Chicago: University of Chicago Press.

Chalmers, David J. (2011). Verbal disputes. *Philosophical Review* 120 (4): 515– 66.

Clark, Andy, and Chalmers, David J. (1998). The extended mind. *Analysis* 58 (1): 7– 19.

Dorr, Cian, and Hawthorne, John (2014). Semantic plasticity and speech reports. *Philosophical Review* 123 (3): 281– 338.

Eklund, Matti (2015). Intuitions, conceptual engineering, and conceptual fixed points. In Chris Daly (ed.), *The Palgrave Handbook of Philosophical Methods*. Palgrave Macmillan, 363-85.

Godfrey-Smith, Peter (1994). A modern history theory of functions. *Noûs* 28: 344-362.

Graham, Peter (2012). Epistemic Entitlement. *Nous* 46 (3):449-482.

Greenough, Patrick (MS) Against conceptual engineering.

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

Leslie, Sarah-Jane (forthcoming). The original sin of cognition: fear, prejudice, and generalization. *Journal of Philosophy*.

Millikan, Ruth (1984). *Language, Thought and Other Biological Categories*. MIT Press.

Monroe, J. Grey, Markman, David W., Beck, Whitney S., Felton, Andrew J., Vahsen, Megan L., Pressler, Yamina (2018). Ecoevolutionary Dynamics of Carbon Cycling in the Anthropocene. *Trends in Ecology & Evolution*; 33 (3): 213 DOI: 10.1016/j.tree.2017.12.006

Neander, K. (1991). Functions as selected effects: The conceptual analyst's defence. *Philosophy of Science*, 58, 168–184.

Plunkett, David, and Sundell, Timothy (2013). Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint* 13 (23): 1– 37.

58 (7– 8): 828– 74.

Scharp, Kevin (2013). *Replacing Truth*. Oxford University Press.

Sosa, Ernest. 1993. Proper Functionalism and Virtue Epistemology. *Nous* 27: 51-65.

Strawson, Peter F. (1963). Carnap' s views on conceptual systems versus natural languages in analytic philosophy. In Paul Arthur Schilpp (ed.), The Philosophy of Rudolf Carnap . Open Court, 503– 18.

Thomasson, Amie (forthcoming). A pragmatic method for conceptual ethics. In Burgess, Cappelen, and Plunkett (eds.), *Conceptual Ethics and Conceptual Engineering*. Oxford University Press.

van Inwagen, Peter (2008). How to think about the problem of free will. *Journal of Ethics* 12 (3/4): 327– 41.

Wright, Larry (1973). Functions. *Philosophical Review*, LXXXII, 139-168.