Conceptual engineering is extremely unlikely to work. So what?*

James Andow†

May 29, 2020

Abstract

Conceptual engineering aims to improve our concepts. That's plausibly an extremely difficult thing to do. Should this make us sceptical of the idea that philosophers should try to do it? You might think so. Cappelen (2018) thinks it shouldn't stop us – but his stated reasons are not really encouraging. In this paper, I say what I think Cappelen should have said, on the basis of a very rough cost-benefit analysis.

1 Introduction

There has been a surge of interest recently in the idea of philosophers trying to improve conceptual resources (see, e.g., Andow, Forthcoming; Cantalamessa, 2019; Cappelen, 2018; Koch, 2018; Nado, 2019; Pinder, 2019; Plunkett and Cappelen, 2020; Prinzing, 2018; Simion, 2018). This interest includes discussion of the ideas that we might do it, that we do do it, that we should do it, and more. There's been interest in how to do it well or how it should be done. The phenomenon of interest isn't always exactly the same but much of recent discussion has been framed around the single label of 'conceptual engineering.'

*This paper is forthcoming in *Inquiry*. This is an author archived pre-print. Please refer to the published article for citations. Thanks to Eugen Fischer and Aimie Hope for helpful comments and discussion. Thanks also to an anonymous reviewer for a different paper of mine whose comments about that paper were the inspiration for this one.

†University of East Anglia

In various places, in this incredibly rich new literature on 'conceptual engineering,' and for various reasons, the idea has raised its head that improving our conceptual resources might be very hard indeed (even if otherwise desirable).[1] The most prominent version of this thought stems from metasemantic concerns. In particular, if metasemantic externalism is true – and it is a very popular position in metasemantics – then conceptual engineering seems like it is going to be very difficult because the control we have over our meanings is severely limited. Such a concern was prominent in Cappelen (2018)'s treatment of conceptual engineering, and here's how Koch (2018, 2) puts it, making clear how the concern serves to cast doubt on how worthwhile conceptual engineering might be:[2]

> On externalism... it seems that changing one's attitudes towards *t* will often not be sufficient to change the meaning of *t*, for *t*'s meaning is at least partly determined by things outside of our cognitive reach. For this reason, it seems that internalists typically allow for a much greater degree of 'meaning control' ... than externalists do. In light of how appealing at least some variants of externalism are to many philosophers, this raises fundamental questions about the activity of CE, such as these: Is our optimism about being able to change significant aspects of our language well founded? If not, then doesn't this support a skeptical stance on [conceptual engineering] as a philosophical method?

While metasemantic issues lie behind the most prominent version of this worry about the difficulty of conceptual engineering, there is a more general form of the concern which might be entertained for other reasons. One might, for example, think conceptual engineering was going to be so difficult as to make it not worth trying simply due to straightforwardly practical barriers – existing conceptual resources are often likely reinforced by existing patterns of thought and behaviour which are things not easily changed.[3]

[1] I put the worry in terms of conceptual engineering being 'very hard' or our 'chances of success being very low.' That's not how it is always put but I think it is the best way to capture the central worry. Here's Cappelen (forthcoming, 11–12) talking about the relevant difficulty in response to recent criticism. "If the meaning of words were easy to control, then English would immediately explode (or implode): there are too many speakers (literally billions, in billions of contexts) with indefinitely many inconsistent preferences, intentions, assessments, goals, plans, and strategies. If English were easy to change, it would collapse. We speakers are fickle, inconsistent and contentious. Our languages are stable and conservative. The latter is in part what makes the former possible."

[2] Plausibly the real challenge is not specifically based on metasemantic externalism but rather the 'inscrutibility' of metasemantics, i.e., the position that "a variety of different factors, some 'internal', some 'external', matter to the determination of a term's semantic meaning and reference, but, in any given case, we don't know precisely which and we don't know precisely how" (Deutsch, 2020, 20). As Deutsch puts it "If metasemantics is inscrutable in this sense, then the intentional effort to render a semantic change in an existing term—an attempted stipulative revision—is just a shot in the dark. Stipulative revisionists can have no clear idea of whether or how or when their stipulations will render the relevant changes" (20)

[3] Indeed that seems to be part of what Cappelen has in mind, see note 1.

To defend conceptual engineering, against the charge that it is too difficult to be worth bothering with, there are various strategies one might adopt. The most straightforward strategy is to attempt to refute the charge. Koch (2018), at a certain level of abstraction, can be seen to adopt this strategy (although he is specifically responding to worries stemming from metasemantic externalism). He argues that it isn't *as* difficult to successfully engineer concepts as you might think (given metasemantic externalism).[4] He argues that:

> For example, in order to make 'woman' apply to the social kind suggested by Haslanger, we have to make it the case that people associate descriptions with the term 'woman' that mark it as a social kind of the envisaged sort. Doing this may not be an easy task, of course, but it is something about which we, as a linguistic community, possess collective long-range control. (19)

And he concludes;

> ...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. Those who thought that changing language was something that could be done by the philosopher alone from the armchair might not be happy to embrace this result—but in my view, we should be careful to avoid this kind of hubris anyway. (21)

How much of an optimistic picture this is—it is certainly much less optimistic about the prospects of armchair-based conceptual engineering than something on a grander scale—is something to be discussed elsewhere. But we can see the general strategy one might adopt: urge for greater optimism about conceptual engineering as a method on the basis of better than assumed chances of success.

A different strategy, for defending conceptual engineering against the charge that it is too difficult to be worth bothering with, is to deny or downplay the supposed link between (a) conceptual engineering being worth bothering with, and (b) conceptual engineering not being extremely unlikely to succeed. This, I take it, is Cappelen (2018)'s own strategy. The slogan forms of Cappelen's view are "The processes involved in conceptual engineering are for the most part inscrutable, and we lack control of them, but nonetheless we will and should keep trying" (72) and "we will keep trying to engage in conceptual engineering and, given the kinds of creatures we are, maybe we must keep trying" (73). But he recognises that this is an odd position on the face of it, and says a few things to

[4]Although, of course, he still thinks it likely pretty difficult. See Fischer (Forthcoming) and Pinder (2019) for other, rather different approaches.

dispel the suspicion that this is a profoundly pessimistic view of conceptual engineering. And it is worth quoting at length (200–201):

> Isn't that all a bit bleak and pessimistic? On the view I defend, the tools we think with are often defective, but there's very little we can do about it.... it's a bit hard to see how I can advocate for and be enthusiastic about such a project. How can this book be a defense of conceptual engineering? ...
> I sympathize with this concern, but I also think that it is deeply misguided. It fails to take into account that what I've just described is an almost universal aspect of large-scale normative reflections. Anyone who spends time thinking and talking about large-scale normative matters should do so without holding out too much hope that their talking and thinking will have significant or predictable effects on the relevant aspect of the world. If you think your views and theories about crime in Baltimore or poverty in Bangladesh will have a significant or predictable effect on either, you're extremely likely to be disappointed... There are of course small-scale local issues where normative reflections will have a direct effect. If I think my daughter shouldn't have an ice cream, then, at least in a few cases, the result will be that she eats no ice cream. Moving to slightly larger-scale issues—say speed bumps in the street where I live—my opinions, views, and pleadings will have tiny effects, but already these effects will be fairly marginal, unsystematic, and unpredictable... On the view proposed ...[conceptual engineering is] far over on the large-scale and unpredictable side. Much closer to crime in Baltimore than to speed bumps in Sofies Gate. ...the worry that I've painted too bleak a picture of the prospects of conceptual engineering simply fails to take into account the relevant comparison class. What I say about conceptual engineering shouldn't be surprising and doesn't make the activity of trying to engineer concepts much different from a wide range of other human efforts to think about how things should be.

There are a few things I want to note about Cappelen's argument here. First, his response to the concern about difficulty relies on successful conceptual engineering being *possible* even if one's chances of success in any given case are vanishingly small (this is also going to be true of my suggested alternative strategy below). However, second, his points about the appropriate comparison class seem to simply miss the mark. Cappelen seems to suggests we compare (A) with (B):

(A) Conceptual engineering done by philosophers as part of big serious well-conceived academic projects;

(B) Ordinary folks going about having speculative dinner table conversations about how to solve inequality or to minor efforts to influence local affairs.

But that can't be right. We *are* tolerant of very low chances of success (I suspect) for (B) but only in the same way that we would be for (C)

(C) Ordinary folks going about having speculative dinner table conversations on the nature of God or to asking a question about Platonism at a public lecture.

Whereas the appropriate comparison for (A) seems to be something more like (D)

(D) Big serious well-conceived policy development projects in institutions such as government departments, think tanks, academic departments such as International Development and Sociology, and NGOs, and well-organised grass-roots campaign groups on local issues.

And it seems much less clear to me that we should be particularly tolerant of (D) where those involved hold out no hope their efforts will have any impact at all. Whatever the reasons for thinking we should be tolerant of small chances of success in the case of conceptual engineering might be, it is not simply that we are always tolerant of small chances of success for idle normative speculation. Such an argument paints a dismal picture that does not do justice to what conceptual engineering can and should be.

A slightly better approach is suggested by the following reflection. What would the expected success rate for conceptual engineering or any of these large scale normative projects have to be for it to be worth trying to bring about such effects? There is some reason to expect a very low rate. Consider that these things only need to be successful once, or maybe once a generation. You certainly don't need conceptual revolution once a week. No proponent of conceptual engineering I know thinks that the population's concept of knowledge should be revised every couple of years. The conceptual engineer dreams only of very occasional revolutions. So, if the expected success rate of a team of engineers is only 1% given five years to work on it, and each five years for fifty years fifty teams worldwide work on the problem, we should assign a very high probability that the object will be realised, and that might mean no further such work is necessary for centuries to come.[5] That suggests it might be appropriate to be content with a very low success rate. However, your suspicion might still be that *this seems like a lot of effort*. In an imaginary scenario such as that we just considered, conceptual engineering does seem very difficult. Can it possibly be worth it? I think that's the key question.

[5]Of course, this assumes that individual conceptual engineers/teams of engineers have a sufficient level of independence in the sense that the chance that at least one of two engineers succeeds, whom each have a 50% chance of success with a problem, is rather greater than 50%. Perfect independence would of course be implausible – as people working on the same stuff talk to each other – but hopefully the level of independence should be expected to be high enough to warrant this back-of-the-envelope calculation.

In the following, my intention is to say what I think Cappelen should have said instead of what he did way. It shares much in spirit and outlook with both Cappelen (2018) and Koch (2018) but differs in strategy. My strategy is not to argue that conceptual engineering is less difficult and so more worth doing than you might think; I will assume it is very difficult (but possible). My strategy is not to downplay a link between the difficult of conceptual engineering and how worth doing it is; I will assume a strong link of this kind. Rather, I just urge that, before writing conceptual engineering off on the basis of difficulty, any critics take into account the *potential value of success* as well as the likelihood of success.[6] In this note, I don't get much beyond some back-of-the-envelope calculations to estimate the expected value of conceptual engineering projects, but they are back-of-the-envelope calculations that seem to help us get some needed perspective on conceptual engineering in philosophy. In §2 I explain the kind of model I'm assuming when making these back-of-the-envelope calculations. In §3 I demonstrate that when we estimate the key parameters the worthwhileness of conceptual engineering turns out to be surprisingly tolerant to extremely small chances of success. §4 wraps up.

2 Some assumptions and a simple model

Before we do some back-of-the-envelope calculations, we first need a quick model to work with and it will help you to see where I'm coming from to articulate a few background assumptions I'm going to make for the sake having a simple model.

1. I concede that conceptual engineering *would* be pointless and shouldn't be attempted if it is impossible.[7] However, I will assume that conceptual engi-

[6] The three basic strategies I mention are not really in competition in the sense that they might be combined in a mutually supportive way. Perhaps the connection between likelihood of success and worthwhileness isn't as direct or strong as might have been supposed but nonetheless exists, the likelihood of success is higher than might have been supposed, and the potential value of success is higher than might have been realised.

[7] In fact, I'm not sure this is true. I have made this assumption elsewhere (Andow, Under review) – and the discussion there is not affected by the qualification I'm about to make – but I do think it is dubious. It is dubious if only because trying to do things which it transpires are impossible can be predictably productive. It is difficult to give an uncontroversial example, but alchemy might be one. I, for one, given the choice to either go back in time to tell everyone lead couldn't be turned to gold (without a particle accelerator) or leave history well alone would leave well alone. But,

neering *is* possible albeit perhaps very difficult (although I won't argue for it here).[8]

2. I will assume that how much of a good idea it is to attempt to engineer concepts is going to be closely related to expected chances of success at least in the sense that it is more worthwhile, ceteris paribus, to attempt to engineer a concept when the chances you will successfully do it are high than when they are low. One simple way to model that is to simply discount any potential pay-off of conceptual engineering according to the risk it won't be achieved if the engineering in question is attempted.[9]

Given this way of thinking about the issue, the concern about conceptual engineering being 'very hard' 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. There are lots of other complications we might consider too, e.g., diminishing marginal gains as more and more philosophers pile on to the conceptual engineering bandwagon, and I'll consider opportunity costs later on. But let's keep it simple for now.

Bearing the above in mind, we can consider a simple model something like this:

Simple model 1 $G_{\mathrm{CE}} = B_{\mathrm{CE}} \cdot P(B_{\mathrm{CE}}) - C_{\mathrm{CE}}$[10]

Where G_ϕ is an indicator of how much of a good idea it is for some time to be spent trying to ϕ (0 being neutral, positive values somewhat of a good idea, negative not); B_ϕ is the benefits of ϕ if successful; $\mathrm{P}(\mathrm{B}_\phi)$ is the

for sake of argument, I'll grant that we shouldn't attempt the impossible.

[8]I take it this is where the real action is: on difficulty rather than possibility. Maybe there are some in principle objections to the very idea of conceptual engineering that are worth taking seriously (the point about metasemantic externalism sometimes seems to be thought of in this way). But I don't buy it. Concepts change; it'd be really weird if it was *impossible* to make it happen deliberately.

[9]For some accessible discussion of discounting on the basis of risk in relation to climate policy, see Caney (2009). He raises some concerns that aren't directly relevant to the discussion here but that might be relevant to a full account of whether and when it is appropriate to expose future generations to the risk of catastrophe and how much we should invest to avoid exposing them to such risks.

[10]Less formally: How much of a good idea it is for some time to be spent trying to conceptually engineer = (Potential benefit of successful conceptual engineering * expectation of success) – (Costs of trying)

probability of success for attempts to ϕ; and C_ϕ is the investment costs of an attempt to ϕ.

G_{CE} is essentially a measure of expected gain. That means the model embodies a particular kind of normative framework with respect to choices in research which I recognise will be controversial.[11] But it is the model I'm going to explore here and I invite others to try other approaches out. So, the next thing to do is to come up with some estimates for those two key value terms: B_{CE} and C_{CE}.

3 Some estimates

First, what are the hoped for benefits of successful conceptual engineering?

At this point, your precise take on conceptual engineering is going to turn out to be very important, as is your background evaluative/normative framework. And I'll make a comment about that at the end. But let's just note that some conceptual engineering projects promise huge returns if successful. A concept of gender that helps the fight against oppression. A concept of marriage that brings security and freedom to hundreds of millions. A concept of corporate responsibility that makes it possible to collect taxes from huge corporations. A concept of global and intergenerational justice that smooths international climate negotiations making effective mitigation possible. And let's adopt the perspective of the would-be conceptual engineer who thinks those things are hugely valuable ends to pursue and is wondering whether to devote time to the project. Should we go ahead?

We need to consider a few things. First, exactly *how* huge are the hoped-for benefits? Can we estimate their value in a way that will allow us to compare them to the costs associated with trying? Obviously, there are various ways we might attempt to force things onto a common metric. You might try to force things onto

[11] In fact, I should note, I think this is a profoundly stupid way to think of the value of research and certainly wouldn't endorse it as metric of the value of research or as a guide in deciding what research to do or fund. However, modelling things in these terms seems to be a helpful way to get a handle on the thought that a research project might be so difficult as to not be worth bothering with. It is also worth noting that, although my examples tend to focus on potential benefits in terms of helping to bring about positive social and environmental change in the world, things like *the intrinsic value of knowledge* might also be among the benefits that might result from a research project.

a direct measure of hedonic utility or adopt some sort of capabilities framework (Sen, 1980). However, since we are playing at cost-benefit analysis let's pretend we think the value of such things can be forced onto the common metric of a dollar value. So we've got a figure to play with, consider the case of climate change. The Stern Review summarised its position as follows, "if we don't act, the overall costs and risks of climate change will be equivalent to losing at least 5% of global GDP each year, now and forever. If a wider range of risks and impacts is considered, the estimates of damage could rise to 20% of GDP or more" (Stern and Stern, 2007, 270). Let's stick with the lower estimate (bearing in mind that though radical at the time the Stern Review itself is now taken to somewhat downplay the problem of climate change).[12] Of course, successfully engineering a concept of justice to help tackle climate inaction is at best one of many things that would be required to ensure we avoid climate inaction, but we can take that into account later on. So just so we've got a number to play with consider that that Global GDP in 2018 was around $85 trillion (The World Bank, 2020) and suppose that we expected zero growth over the next 50 years, and that we only need to consider the next 50 years as beyond that there are no costs of climate inaction. That's around $4 quadrillion, and a GDP loss of around 5% of that would be around $200 trillion. And we've underestimated those costs at every step of our back-of-the-envelope calculations. But nonetheless, let's use it as an estimate of the costs of climate inaction that we might hope our conceptual engineering will help avoid. So the potential benefits are large. That suggests we might be able to be tolerant of low chances of success. But of course there will come a point where, taking into account the costs of trying to engineer a concept, the chances of success are so low that it is not worth trying even if the potential pay-off is extremely high: there'll be a break-even point. To get a sense of where it might be, we need to think about how much conceptual engineering costs.

How much does conceptual engineering cost? Conscientious conceptual en-

[12]As Stern himself said in 2013, "Looking back, I underestimated the risks. The planet and the atmosphere seem to be absorbing less carbon than we expected, and emissions are rising pretty strongly. Some of the effects are coming through more quickly than we thought then" (Stewart and Elliott, 2013). And it is a long way from clear that the kind of exercise the Stern Review was carrying out comes anywhere near capturing the full extent of the losses that a climate catastrophe would represent to the world.

gineering might well be rather more expensive than some other types of philosophy. Certainly, to be fair to anyone who suspected conceptual engineering might be prohibitively costly, we should expect the costs of a conceptual engineering project that even remotely aspires to have the kind of global influence necessary to help avoid the costs of climate inaction to be pretty high. It might require serious empirical work, perhaps, and stakeholder engagement, certainly some sort of dissemination strategy, perhaps it will require hired marketing consultants and lobbyists, perhaps even hiring lawyers, paying influencers, or commissioning Netflix series – depending on the project. For a ballpark overestimate, let's look at what a big project in social science research looks like. The UK's Economic and Social Research Council's research grants fund up to £1 million (say $1.25 million) per project, but let's go wild and suppose you would need *forty* times that to develop a meaningful conceptual engineering project (perhaps including huge quantities of follow on funding to help facilitate 'impact'): £40 million (say $50 million).[13] That's around one four-millionth of our sketchy radical underestimate for Global GDP over the next 50 years.

So, what's the break-even point in terms of chances of success given these estimates? A project that potentially solves climate inaction would be worth pursuing (putting aside opportunity costs) just so long as the expectation of success was greater than around 1/4,000,000. And even if we adjust this to consider the fact that a conceptual engineering project would at best be a part of more extensive effort – it is going to remain a pretty low number. Perhaps, we should suppose that a radical transformation of global concepts of justice is only 1 part in 4,000 parts of the solution to climate inaction, i.e., that we should adjust the break-even point to be as high as an expected success rate of 1/1,000.

We do need to factor in opportunity costs, at some point. But let's not complicate things by asking about opportunity costs associated with conceptual engineering *as opposed to training as a virologist or human rights expert* and confine the opportunity costs we consider to those relating to other philosophical activities

[13]For context, that's around 20% of the annual WWF-US budget for both 'Conservation field and policy programs' and 'Public education' (World Wildlife Fund US, 2019), and also around 20% of the annual budget of the ESRC (Department for Business, Energy, & Industrial Strategy, 2018).

(but let's not pretend that if the \$50 million doesn't go to conceptual engineering projects run by philosophers then it would go to philosophers to do something else). What philosophy as a profession has to spend in various ways is *time philosophising*. So the idea is that we assume that we know how much philosophy is to be done but are unclear as to whether some amount of it should be conceptual engineering, and now we're going to think about whether the prospective benefits of it being conceptual engineering are worth the costs.[14] Taking these into account we know how to think about it using the following model.

Simple model 2 $G_{\text{CE}} = B_{\text{CE}} \cdot P(B_{\text{CE}}) - C_{\text{CE}} - B_{\text{O}} \cdot P(B_{\text{O}}) + C_{\text{O}}$[15]

Where the subscript 'O' concerns 'doing something other than CE.'

So let's think about how high those opportunity costs would need to be to make a difference to our verdict of whether it is a good idea to attempt some conceptual engineering.

Let's be super charitable to the alternative option. Suppose that doing something else was cost-free somehow. And suppose that it was somehow certain to bring about its benefit. What happens to the success rate required for conceptual engineering to break even as the potential pay-off for the alternative project rises? The answer is that the break-even point remains below 1% until the expected pay-off of the alternative project reaches \$450 million. Now I don't know how to quantify the value of the average successful project in pure conceptual analysis or theoretical philosophy or whatever—whether we focus on the intrinsic value of the knowledge produced, the benefits to society, or both—but that looks quite high to me and it would have to be higher to reflect any costs or risk associated with the relevant activities. In any case, the point of all this isn't to show that conceptual engineering will always trump doing something else within philosophy.

[14] I assume here a sharp division, and one transparent to the philosophical community, between conceptual engineering projects and other projects in philosophy. Maybe that's implausible (as is, in fact, implicit in my discussion in Andow Forthcoming). And it is worth thinking through the implications of that for these kinds of back-of-the-envelope calculations in general. However, for the kinds of big projects that are deliberately and explicitly engaged in conceptual engineering that are under consideration here, it's not so implausible to assume a sharp distinction.

[15] Less formally: How much of a good idea it is for some time to be spent trying to engineer some concepts = (Potential benefit of successful conceptual engineering * expectation of success) - (Costs of trying) - (Missed potential benefit of doing something else instead * expectation of success) + (The avoided costs associated with doing something else)

The point is that it looks a competitive option even if the chances of success are extremely small.

How do things look if we replace the 'big money' project with five years in a career of armchair musings about concepts and how they should be? Obviously, since the costs are rather lower (5 years at 40% research time of a starting professor salary at my institution would be around \$150,000), the project is going to be even more worth pursuing if both the potential pay-off and the chances of success were to be kept constant. However, while keeping the potential pay-off constant makes sense, given our question, keeping the chances of success constant does not. As both Cappelen and Koch seem to agree, our expectation of the success of armchair musings should be very low indeed. A conscientious and well-resourced project plausibly has chances of success orders of magnitude higher than armchair musings. Nonetheless, were we to use the same estimates for the potential pay-off (and the part of conceptual engineering in bringing about that pay-off), the break-even point, not taking into account opportunity costs, gives a chance of success of 3 in 1 million (which maybe doesn't seem ridiculously implausibly high). Adjusting this break-even point to accommodate various levels of opportunity cost until the break-even point reaches something still plausible, e.g., orders of magnitude below the 1% figure used above. At a rate of success of 0.1% the project only stops being worthwhile when the potential pay-off of the alternative option rises to around \$50 million. At a rate of success of 0.01% the project only stops being worthwhile if the potential pay-off of the alternative rises to around \$5 million. Although, remember that those are treating the pay-off of the alternative as being guaranteed if attempted. Nonetheless, I suppose it isn't obvious armchair conceptual engineering isn't going to be a competitive option – your view on this is going to depend on how you value typical projects in philosophy. Although, if big project money could really buy you one-hundred times the chances of success it would seem worthwhile investing.

4 Conclusion

Now, I understand this is all just back-of-the-envelope stuff. It is playing with a toy example and a very simplistic model. But it serves to make a point. I think the significance of the *difficulty* (for whatever reason) of deliberate conceptual change is overestimated in the conceptual engineering literature – especially given the huge potential pay-offs conceptual engineers could be working to pull off in many cases.

You will have noted a fair way back that I have focused on a very specific kind of conceptual engineering: that aiming to bring about large scale change that will benefit huge numbers of people. The same points will not follow for conceptual engineering projects with only very minor, short-term, projected pay-offs even in a best-case scenario. But the back-of-the-envelope calculations are worth doing for yourself for projects with ambitions that involve re-engineering society's big concepts – honesty, truth, fidelity, authenticity, freedom, responsibility, and so on. It is going to be more difficult to put anything like a dollar value on the kind of transformations that might be achieved, or on the intrinsic value of having better versions of these concepts. And maybe they are not going to be on a par with saving the planet from climate change. But, of course, we made adjustments when thinking about that case for the fact that conceptual engineering would only be a tiny part of the solution. Whereas, conceptual engineering might be all that is needed to set in motion a transformation of such 'big concepts'. My guess is that we don't need much expectation of success for it to be worth giving conceptual engineering projects of this kind a shot. Nonetheless, perhaps there is an argument in the above—although I don't intend that to be the take-home point of this short note which is just to encourage people in this literature to do the back-of-the-envelope calculations and to note that having done a few myself it seems conceptual engineering might be worthwhile even if there is little chance of success—that insofar as philosophers are going to engage in conceptual engineering, they'd do well to tackle the big shit and to do it seriously.

You will also have noted that I haven't attempted even a back-of-the-envelope full cost-benefit analysis and, notably, have not considered the prospect that an attempt to engineer a concept might have unplanned negative impacts. That's something which deserves to be taken into account in a full analysis.[16] Presumably there would always be a risk that your attempt to meddle in concepts for the better in fact changes things dramatically for the worse. For example, maybe doing so aiming to promote climate change action could end up fuelling a popular backlash that results in the collapse of existing climate change mitigation measures and climate catastrophe. Take that into account and maybe the full cost benefit analysis is going to radically alter. However, if that's right, that's a subtly different objection to conceptual engineering than the one under consideration here (although it is potentially connected, e.g., through the link with metasemantics and the extent to which meaning change is 'out of control').

In sum, I'm urging for optimism about conceptual engineering on the basis that it seems the prospective pay-offs might more than make up for the small chances of success. But that comes with caveats. The optimism I am urging is for serious large-scale projects with associated greater chances of success that involve a prospect of very large pay-offs and with no significant risk of serious negative consequences. The smaller and less serious the effort, the smaller the prospective pay-off, and the more worrying the possible unintended consequences, the less optimism about conceptual engineering this kind of reasoning is going to warrant.

References

Andow, J. (Forthcoming). Intuitions about counterfactual cases as evidence (for how we should think). *Inquiry*.

Andow, J. (Under review). Fully experimental conceptual engineering. *Under review*.

Caney, S. (2009). Climate change and the future: Discounting for time, wealth, and risk. *Journal of Social Philosophy 40*(2), 163–186.

[16] A full analysis might also want to consider the question of whether it is appropriate to discount future costs and benefits in some of these cases. There at least seems to be something to be said for getting a better concept of X sooner rather than later, and it is worth thinking about how to model that appropriately.

Cantalamessa, E. A. (2019). Disability studies, conceptual engineering, and conceptual activism. *Inquiry*, 1–30.

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

Cappelen, H. (forthcoming). Conceptual engineering, topics, metasemantics, and lack of control: Responses to sarah sawyer, laura schroeter, francoisschroeter, and tim sundell. *Canadian Journal of Philosophy*.

Department for Business, Energy, & Industrial Strategy (2018). *The Allocation of Funding for Research and Innovation*.

Deutsch, M. (2020). Speaker's reference, stipulation, and a dilemma for conceptual engineers. *Philosophical Studies*.

Fischer, E. (Forthcoming). Conceptual control: On the feasibility of conceptual engineering. *Inquiry*.

Koch, S. (2018). The externalist challenge to conceptual engineering. *Synthese Online first*.

Nado, J. (2019). Conceptual engineering via experimental philosophy. *Inquiry Online first*, 1–21.

Pinder, M. (2019). Conceptual engineering, metasemantic externalism and speaker-meaning. *Mind Online first*.

Plunkett, D. and H. Cappelen (2020). A guided tour of conceptual engineering and conceptual ethics. In H. Cappelen, D. Plunkett, and A. Burgess (Eds.), *Conceptual Engineering and Conceptual Ethics*. Oxford University Press.

Prinzing, M. (2018). The revisionist's rubric: conceptual engineering and the discontinuity objection. *Inquiry 61*(8), 854–880.

Sen, A. (1980). Equality of what? *The Tanner lecture on human values 1*, 197–220.

Simion, M. (2018, November). The 'should' in conceptual engineering. *Inquiry 61*(8), 914–928.

Stern, N. and N. H. Stern (2007). *The economics of climate change: the Stern review*. cambridge University press.

Stewart, H. and L. Elliott (2013). Nicholas Stern: 'I got it wrong on climate change – it's far, far worse'. *The Guardian*.

The World Bank (2020). *Structure of Output (Table 4.2)*.

World Wildlife Fund US (2019). *2019 WWF-US Annual Report*.