

What is Conceptual Engineering and What Should it Be?*

David J. Chalmers

I was going to call this paper “Conceptually Engineering Conceptual Engineering” but I found that that title was already taken by an excellent article by Manuel Gustavo Isaac. Instead I’ve chosen a title paying tribute to the Sally Haslanger mode of conceptual engineering (“Gender and Race: What are they? What do we want them to be?”). My original thought was to give a talk with a bunch of theses about conceptual engineering, where the only case study is engineering the concept of conceptual engineering. That got too vertiginous, but there will still be some of that. I’ll start by talking about the concept of conceptual engineering and distinguishing different varieties, and go on to issues about the importance and the difficulty.

What is conceptual engineering?

What is conceptual engineering? There is an obvious way to come at this. To find the definition of conceptual engineering, go look up the definition of engineering, and then just appeal to compositionality, applying the definition to concepts wherever possible. There are many definitions of engineering, often put forward by different engineering associations such as the mechanical engineers, the civil engineers, and the electronic engineers. But these definitions have a common core. There’s a fairly common simple definition, which a lot of the more complicated definitions are basically variations on.

On this definition, engineering is the process of utilizing knowledge and principles to design, build, and analyze objects. The key thing there is design, build, and analyze. You’ll find variations on those three things in most definitions of engineering. Some of them extend the list to about twenty different things. There’s to operate, and to maintain, and to repair, and to forecast, to evaluate, and so on. But they mostly fit within this rubric of designing, building, and analyzing.

⁰This is a lightly edited transcript of a talk given at the NYU workshop on “Foundations of Conceptual Engineering” on September 14, 2018. Thanks to Banafsheh Beizaei for the transcription, to Vera Flocke for organizing the conference, and to the audience for discussion. References to be added later!

Invoking compositionality, conceptual engineering will be something like the process of designing, building, and analyzing concepts. That's not a bad definition, except that 'analyze concepts' already has a meaning which is not totally apropos in this context. Maybe 'evaluate concepts' is better. And maybe 'implementing' is better than 'building' where concepts are concerned. With these tweaks, we get the following definition: conceptual engineering is the process of designing, implementing, and evaluating concepts. That is not a bad first pass at what conceptual engineering is all about.

This definition gives us different broad stages or types of conceptual engineering. There's the design stage, where we design concepts. There are various ways to do that. One classic way would be to give a definition, or maybe an inferential role, or some paradigm cases, or something like that. Next is the implementation stage, like the stage where you actually build the bridge. In the implementation stage you actually use a concept, and perhaps try to get others to use it too. This is what Herman Cappelen calls conceptual activism. And then there's the key evaluation stage, which plays a central role in the conceptual ethics work by people like Alexis Burgess and David Plunkett. Here what's key is the evaluation of how good these concepts are in themselves and for certain purposes, to see how well they play key roles.

You can see all three of these things playing a role, say, in bridge engineering. You design a bridge, you implement a bridge, you evaluate the bridge to see how well it's doing. If the evaluation isn't positive, you design some repairs and you implement the repairs. And so on. You also see something like this in software engineering. You design a program, implement the program, evaluate the program, and so on in a continuing circle.

Incidentally, I like the software engineering analogy because many people worry about how conceptual engineering can work if concepts are abstract objects. It's not obvious that you can build a new resident of the third realm in this way. Exactly the same issues arise for software engineering. Programs are usually thought of as abstract objects, and there are complicated things but fairly obvious things you can say in both cases about what goes into the implementation stage. If programs and concepts are pre-existing abstract objects, then what's relevant to implementing a new program or a new concept is a matter of coming to stand in new relations to those things. The program comes to be realized on a computer, and the concept comes to be grasped in people's psychology. Anyway, I'm not going to make heavy weather of abstract objects versus concreta here.

A more important issue here is a distinction between creating versus fixing as kinds of engineering. In creating a bridge, we design and implement a new bridge from scratch, perhaps after a

period of evaluating what is needed. In fixing a bridge, we evaluate an old bridge and then design and implement repairs. This is very relevant in ordinary engineering, very relevant in software engineering, and ought to be very relevant in conceptual engineering. This distinction will be what I'm focusing on to a considerable extent.

Examples of conceptual engineering

First I want to just go through a bunch of examples of what I think of as paradigmatic conceptual engineering, in both modes: the creating and the fixing mode. These examples are all within philosophy. That's not because philosophy is the only locus of conceptual engineering. I think it's absolutely everywhere. But philosophy is the one that I know best and that many of us here have particular expertise on. So we might as well talk about what we know.

Take something from metaphysics. For me, the concept of supervenience is a paradigm example of conceptual engineering. Someone said it has the smell of something that was thought up in the metaphysics lab. This idea of these properties, which depend on some other class of properties in that you duplicate one, you duplicate the other. This concept, I think, was engineered over the twentieth century—Moore had it without the name, Hare introduced the name, Davidson and Kim and others made much of it. To me that's paradigmatic conceptual engineering. And indeed the notion of supervenience was once thought to be one that could do a lot of philosophical work that previous concepts like identity might have been hoped to do as well. Then later on people thought, okay, that doesn't do that work quite so well, so now we've got these newfangled concepts like grounding. This involves a bit more conceptual engineering or at least conceptual abstraction of ideas which are in the air, under a new label. So these are examples in the mode of creation. In the mode of fixing I think of Amie Thomasson's work on existence in metaphysics, which takes a pragmatic approach to which notion of existence or of object will serve us best.

I think conceptual engineering in the mode of creation is everywhere throughout the philosophy of language. Semantic values and notions of meaning are engineered all the time. Carnap's notion of intension, Frege's notion of sense, Grice's notion of implicature, Kripke's notion of rigid designator. I think all of those can be seen as engineering fruitful new concepts in the philosophy of language. In the fixing mode, there's various work on truth. Actually Carnap's work on explication took truth as one of his central examples, and he saw Tarski's explication as a kind of conceptual engineering. More recently Kevin Scharp has diagnosed truth as an inconsistent concept, proposing his upward and downward replacements.

In the philosophy of mind, my colleague Ned Block is a paradigmatic conceptual engineer. His 1995 paper on the function of consciousness says consciousness is a mongrel concept, problematic in all these ways, let me find you these more precise and interesting concepts, such as the concept of access consciousness, engineered and useful. Herman in his book on conceptual engineering puts forward my paper with Andy Clark on the extended mind as a paradigm case of conceptual engineering, in engineering the concept of belief. My attitude towards that is slightly complicated, and I'll get back to that shortly. Belief has actually been one of the central candidates for conceptual engineering since the 40s and 50s, with Carnap and others giving it various explications. More recently Tamar Gendler has been engineering a new concept in the vicinity of belief with her concept of alief.

In social philosophy you find this kind of thing all the time, I think. People pointing to phenomena that may have been overlooked with useful concepts, or drawing distinctive concepts out of strands of discussion. So I think Miranda Fricker's work on epistemic injustice and its varieties like testimonial and hermeneutic injustice, would be a paradigmatic example here of drawing out a fruitful concept. Sally Haslanger's work on gender and race is another. The paradigm example would be her work towards the analysis of the concept of woman in terms of oppression. What Sally calls ameliorative analysis is conceptual engineering in the revisionary mode to serve various ends, including the ends of social justice. This ameliorative strand of conceptual engineering has been picked up by many other people in recent social philosophy. Kate Manne's revisionary analysis of misogyny is an example.

In metaphilosophy, which is roughly the field of this conference, you can find a few paradigmatic examples. I think Carnap on explication was itself a wonderful example of relatively de novo conceptual engineering. Here is a new and useful concept. And indeed the very concept of conceptual engineering (you knew it was coming) is a marvelous piece of conceptual engineering. By the way, I've noticed in the recent literature on conceptual engineering, the credit for the phrase 'conceptual engineering' often goes to Simon Blackburn in 1999. The credit should really go to Richard Creath in 1990. Creath made a big deal of Carnap as a conceptual engineer. It first plays a role in his 1990 book *Dear Carnap, Dear Van*, and many other papers. Conceptual engineering has been all over the Carnap literature for decades. PhD theses have been written on it. As Manuel Gustavo pointed out to me, Carnap himself actually talks about linguistic engineering in a couple of places, but never uses the phrase conceptual engineering. In any case this concept of conceptual engineering is itself a model of a useful bit of conceptual engineering.

De novo conceptual engineering vs. conceptual re-engineering

So now with those examples in place, let's think about the whole fixing versus creating distinction. Herman's definition of conceptual engineering basically ties it very centrally to the fixing project. The title of his book is *Fixing Language*. I think the more detailed definition is assessing and improving our deficient representational devices. I'd say that's part of conceptual engineering. If we said that civil engineering, say bridge engineering, was the project of fixing and improving our deficient bridges, or that software engineering was fixing and improving our deficient programs—well, that's part of software engineering and bridge engineering. But, between us, it's not necessarily the most important or most exciting part. There's also the whole project of building new bridges and building new software. I think Herman in his book acknowledges the possibility of this project, and goes on to say something like, that's just not the one I'm interested in here. That's fine, but I do think it's worthy of our attention.

So I encourage making a distinction between what I call de novo engineering and re-engineering. De novo engineering is building a new bridge, program, concept, whatever. Re-engineering is fixing or replacing an old bridge, program, concept, or whatever. The name is still up for grabs. At one point I was using de novo versus de vetero, but someone pointed out to me that wasn't really proper Latin. It's not totally straightforward to draw the distinction. There are some hard cases. Take the Tappan Zee Bridge, just up the Hudson River from New York City. The old Tappan Zee bridge is still there, and they're building a new bridge in the same location as the old bridge, in order to replace the old bridge. Is that de novo engineering because it's a new bridge, or is it re-engineering because it's a replacement? For my purposes I'm going to count that kind of thing as re-engineering, because the central theme is it's being used to somehow to fix an old bridge. Likewise in conceptual re-engineering, the point is to fix this very concept. We can argue about how we draw the lines.

Many or most of the standard examples in the conceptual engineering literature are conceptual re-engineering. Certainly the Carnapian explication literature is very much a literature on re-engineering. Georg Brun has a nice paper on explication as conceptual re-engineering—he uses that exact phrase. The conceptual engineering of belief has largely been re-engineering. There's also been re-engineering with the concept of truth. More recently there has been a lot of interest in re-engineering social concepts such as the concept of woman and the concept of race.

Many of the examples I gave, on the other hand, I think, look more like de novo conceptual engineering. Epistemic injustice, supervenience, rigid designator, and, indeed, conceptual engi-

neering. These weren't particularly trying to fix or replace other concepts. If you squint really hard, you might supervenience is intended as a replacement for identity. But that's not quite right. The concept of identity is doing fine. It's just that there's a job people were using identity for, in some reductive projects, that people then tried to use supervenience to do. Maybe you could argue that the concept of conceptual engineering is a replacement for the concept of explication. I think it's not the most productive way to think about these concepts as re-engineerings. Rather the spirit is *de novo*, building rather than fixing. Here are some useful, fruitful concepts that we can use to do some interesting philosophy with, and that may have useful consequences.

My proposal is that conceptual engineering either includes or should include—that's a little bit of weaseling—both *de novo* conceptual engineering and conceptual re-engineering. The “should” version may seem to suggest that conceptual engineering doesn't already cover both, in which case this would be a proposal for conceptually re-engineering conceptual engineering. If so, well, conceptual re-engineering is important too, and this would be an example of it. But my own view is that conceptual engineering already covers both. This is partly in virtue of compositionality. It'd be very weird if conceptual engineering worked in such a different way from other kinds of engineering. It's partly in virtue of the unity of the category and partly in virtue of other factors. I may be wrong. Maybe the use of conceptual engineering by various people in this community recently focusing on re-engineering has given us some semantic glue that makes it stick to that. So I don't want to put too much weight on the semantic claim. That's really a verbal dispute about conceptual engineering.

This leads to my sideline on conceptually engineering belief. What I said about my attitudes towards the conceptual engineering of conceptual engineering was Andy Clark's and my attitude to the conceptual engineering of belief back in “The Extended Mind”. Herman quotes us saying: “We don't intend to debate what is standard usage; our broader point is that the notion of belief ought to be used so that Otto qualifies as having the belief in question.” Here Otto is the guy who carries around the notebook which his memories are stored in. I'm happy to be taken as a paradigm conceptual engineer, but I'm not sure that we saw this as conceptual engineering. Our own view was that these extended cases of beliefs were literally beliefs. So the word ‘belief’ already covers them, perhaps in virtue of unity and the explanatory fecundity of the category. We were just saying, if someone wants to argue with that, here is the underlying more important claim.

I'm inclined to say something like this about about the meaning of conceptual engineering. I think it does cover *de novo* conceptual engineering. If you think that's not the actual meaning of the term, then I'll argue that it *should* be used to cover *de novo* conceptual engineering. My

metaphilosophical ideology is that you shouldn't get too hung up on words. On the other hand, words matter at least for practical purposes. This was roughly our attitude in "The Extended Mind". Andy and I could have introduced a new term, 'e-believe', to cover all these extended cases, and made claims about how unified e-belief is with the ordinary cases of believing and how e-belief plays the most important role. We could have done that, but what fun would that have been? The word 'belief' is used a lot, it's got certain attractions in explanation, so attaching the word 'belief' to a concept plays certain pragmatically useful roles. Likewise the word 'conceptual engineering'. Conceptual engineering is cool, people have conferences on it. So using that name has what Herman calls a lexical effect. Using only another name, like de novo this and that, is not going to work in the same way. So pragmatically it makes sense to try and attach this thing you're interested in to this word. And I think doing that yields a deeper and richer subject matter.

I know some people want to restrict the term "conceptual engineering" to conceptual re-engineering. Of course I do allow that there is a legitimate narrower subject matter of conceptual re-engineering with some distinctive concerns. At the same time, words have power, and "conceptual engineering" is a powerful expression. If people had just talked of "concept repair" or "concept re-engineering" from the start, I don't think the interest would have been so great. The power comes in part because of the suggestion of something that goes beyond mere fixing. I'm trying to vindicate some of that power.

Is this a verbal dispute over conceptual engineering? Yes, but there's an underlying nonverbal point. Conceptual engineering should cover de novo conceptual engineering, because of unity with conceptual re-engineering and because de novo conceptual engineering is at least as important in philosophy and elsewhere as conceptual re-engineering.

Homonymous and heteronomous conceptual engineering

This immediately leads us to delicate questions about the connection between conceptual engineering and linguistic engineering. Sometimes these phrases are used almost interchangeably. Certainly it's the case that wherever you find conceptual engineering, you almost always find some linguistic engineering. Whenever there's a proposal about a new concept, there's also a proposal about a word for it to be attached to. This is a version of what Peirce calls "the ethics of terminology", which Steve Yablo turned me on to years ago. All this really kind of comes down to the ethics of terminology, or the ethics of language—how we should be using new and old words to express new and old concepts.

In the linguistic mode, there's a related distinction between homonymous and heteronymous conceptual engineering. Homonymous is same-word linguistic engineering—fixing a new meaning for an old expression. Cases like 'woman' or 'misogyny' and 'truth' are homonymous. The central cases of homonymous conceptual engineering are cases of conceptual re-engineering: fixing the old concept expressed by the word. There will be some non-central cases where you can use an old word for something totally new: the word 'supervenience did' already exist, with a different meaning. I don't think it's usefully regarded as fixing the meaning of that word. In that case it's more or less coincidental that that word is being used. Heteronymous conceptual engineering is different-word linguistic engineering. For example, Ned re-engineered the concept consciousness to the concept of access consciousness. Instead of saying, let's use the old word to express access consciousness—that would be homonymous—he said let's introduce a prefix. access consciousness, to express the new concept. Using a prefix is a very very common strategy in philosophy, and that makes it heteronymous. In other cases there's a new concept and a new word: supervenience and, indeed, something like conceptual engineering may be like this.

The existing literature focuses almost entirely on the homonymous case, but I think the heteronymous case is just as interesting. One thing that makes it interesting in this context is that it's at least somewhat less vulnerable to challenges such as Strawson's challenge: aren't you just changing the topic? To which the answer for heteronymous conceptual engineering is, sure, I'm changing the topic—here's an interesting topic. It also helps with the externalist challenge. It's not that that worries about externalism don't arise in coining a new word. But in the use of an old word, they arise big time, because there's that whole social community of users of the word to defer to, making it all the harder to effect conceptual change.

You might ask why the homonymous/heteronymous distinction matters. When it comes to linguistic engineering, I'm on record in my paper on verbal disputes as saying that the words we use for things don't matter all that much in philosophy. So why does it matter whether we use a new word or an old word to make our claim? What's the point of homonymous conceptual engineering or indeed of heteronymous conceptual engineering? Here I think we've got to distinguish here between broadly theoretical and broadly practical projects. For theoretical purposes, I think in principle anything you can say with homonymous conceptual engineering can also be said with heteronymous conceptual engineering, and vice-versa. You can use an old word, you can use a new word. For theoretical purposes it's just a verbal difference between those two. And indeed, as I say in the verbal disputes paper, for an ideal reasoner this difference wouldn't matter. But we aren't ideal reasoners. We have strong associations with the use of certain words. So it makes a

big practical difference whether one uses a new word or an old one.

Of course projects in heteronymous conceptual engineering could be done homonymously, and vice versa. Instead of introducing the new word ‘supervenience’, we could say, here is what ‘reduction’ should mean. And then give a definition of reduction in terms of supervenience. Or instead of using the word ‘sense’ let me appropriate the word, let’s say, ‘meaning’. Things could be done that way. And vice-versa for various homonymous projects, like the revisionary analysis of truth. I think Kevin Scharp does end up using a different phrase, ‘upward truth’. For Sally’s ameliorative analysis of ‘woman’, there’s certainly going to be a related heteronymous project that says, let’s use a new word, like ‘womyn’, to express the new concept. But the practical effects will be different. The upsides and downsides here are mainly practical rather than theoretical, although the lines can blur in some cases. Sometimes we can achieve our purposes better with a new word, and sometimes with an old word. I think tomorrow Ari Koslow is giving a paper on the upsides and downsides of using the same words and different words.

Just a couple of initial thoughts about this. There’s entrenchment: new words are expensive and harder to get people to use. Old words often have associated prestige. ‘Conceptual engineering’ is prestigious—people are having a conference on it! So if you want to get your concept adopted, attach it to that powerful engine. There’s also roles. An old word has certain fixed roles. Homonymous conceptual engineering can allow a new concept to very easily become associated with those roles. Certainly, I think, in the case of, say, gender concepts like woman, look at all the roles that plays in society; you want something different to play those roles, and this is the way to do it. This ties closely to projects in social justice. Given fixed social roles for terms like ‘woman’ or ‘marriage’, homonymous conceptual engineering allowing a new concept to play those social roles, can in many cases make for a more just world. So those are some of the upsides.

There are downsides too. Homonymous conceptual engineering, especially for theoretical purposes, can be very confusing, with all these multiple meanings floating around. It can also be extremely difficult to implement—that’s something which Herman has focused on. Unless one is very powerful or very lucky, or in a small community. There’s only thirty people in the world working on conceptual engineering and they’re all in the room, so you have a chance. But even then it’s not a big chance. Heteronymous conceptual engineering is perhaps somewhat easier to implement, at least in limited circles. It’s not so hard to put forward a new word, though still not easy to get it adopted. Those are upsides and downsides.

Conceptual engineering and thesis engineering

Let me say something about the importance of conceptual engineering. Some people say it's all of philosophy. Is it important? Yes, some of the most important advances in philosophy have quite clearly involved conceptual engineering. That goes especially for some of the cases I gave of de novo conceptual engineering—rigid designation, implicature, epistemic injustice, and so on. You're making new concepts to do a lot of work.

Is conceptual re-engineering important in philosophy? It's certainly practically important, especially for practical roles in the wider world, such as achieving social justice. Philosophers have an important role to play in that project—both the activist part and the more theoretical part of figuring out which concepts can best play these roles. Aside from these roles in the wider world, conceptual re-engineering can also be theoretically useful within the philosophical community, for example cleaning up a concept, modifying a concept so it can better play explanatory roles, and distinguishing different concepts that play different roles. Maybe the different notions of probability and consciousness provide examples of that. I think most of the time this can be done just as well in heteronymous or de novo mode. We'll come back to the reasons for that in just a moment.

Is conceptual engineering everything in philosophy? This is clearly false. Lots of important philosophy involves arguing for theses using old language, and is none the worse for that. Think case like Jackson's knowledge argument or Parfit's repugnant conclusion. As far as I can tell, there's no real conceptual engineering in these cases. You can use the old concepts still for very interesting phenomena and theses.

Even when philosophy involves new concepts, I think it typically also involves new theses involving those concepts. I think the role of theses here is crucial. Why is supervenience interesting? You don't just look at supervenience and say, that's a cool concept. Rather, it serves some potential roles for you that can connect to certain theses. Many people think, for example, that physicalism requires supervenience. Some people think if you have supervenience, then you get physicalism. Supervenience suffices for various reductive projects. Then someone like, say, Kit Fine comes along and argues with that, and we can assess those theses.

Or take implicatures. The concept of implicature is interesting, but really where the rubber meets the road is theses about where implicature happens, or about which maxims govern implicatures. Implicatures play these roles in discursive practice, and so on.

In general theses are where the rubber meets the road. Theses have a certain primacy in

philosophy, at least the sort of philosophy aimed at discovering truths. In this sort of philosophy, conceptual engineering is largely driven by thesis engineering. The debate over grounding and supervenience, which concept is best in metaphysics, really comes down to what are they useful for? What theses can they be used in, for example, for reductive projects like physicalism, is grounding or supervenience the best conditional. For assessing the status of various concepts of meaning, let's look at the different properties of language that they can explain, theses you can get them involved in.

This is not saying that concepts are not important in truth-aiming philosophy. It's to say that the importance of concepts derives from the importance of theses that they're involved in. This is something that Carnap tacitly recognized in his discussion of explication, where he talked about the fruitfulness of a concept as being on one of the dimensions of assessment. The fruitfulness of a concept comes down to the number of interesting and useful and explanatory theses it might be involved in. So on this view the importance of concepts derive from the importance of theses, and certainly formulating theses and arguing for them goes well beyond conceptual engineering.

It's true that philosophy has many aims distinct from discovering truths. Philosophy also aims to raise questions, to help us understand, to help us see the world differently, to live better lives, to improve the world, and so on. New and improved concepts certainly can help with those things. For example, once you have the concept of epistemic injustice, you see all sorts of old situations in a new way, and this can help achieve more just outcomes.

Even in these cases I don't think a new and improved concept has non-derivative importance. What's important is the role it plays, in this case in helping us see the world in a new way and in improving the world. I think that very often these roles are tied to theses— such-and-such cases are cases of epistemic injustice, epistemic justice is bad in a certain distinctive way, and so on. Where practical projects are concerned, normative and evaluative theses will often be central to the philosophical grounding of these projects. But perhaps aiming for true theses doesn't capture all that's going on. So perhaps more generally, one should speak of the primacy of roles rather than the primacy of theses, and understand the fruitfulness of a concept more generally as tied to the roles it can play. And more specifically, at least a very common mode of analytic philosophy is mainly aimed at discovering truths, and at least in this mode thesis engineering has priority over conceptual engineering.

Conceptual engineering and concept pluralism

This leads to the metaphilosophical view of these things which I favor, concept pluralism, which Herman also cites in this book as a kind of conceptual engineering view. There are many concepts in the vicinity of most philosophical words playing different roles. We ought to articulate those roles and find the concepts that play them best. I think that's a kind of conceptual engineering: finding concepts that play these roles the best. We're evaluating concepts in many cases, finding relatively new concepts that can play them. That's a kind of conceptual engineering. For the most part I don't care whether it's homonymous or heteronymous, except for these practical purposes. But again, it's not just conceptual engineering. It's articulating theses of the form 'x plays role y'. You get a cool concept when you get some cool theses.

For a concept pluralist, de novo conceptual engineering is often better than re-engineering. This is the strongest thing I'm going to say in favor of de novo conceptual engineering. My paradigm of conceptual re-engineering is that you somehow fix or replace the old concept, so the old concept is no longer around. For the concept pluralist, a lot of time you want to say, why not have both? Even, say, for the case of Sally's analysis of woman. Even if you think that that the old biologically-based concept of woman was an unjust concept for various social purposes, somebody might still think, well, nonetheless it's a useful concept to have around, say for certain medical purposes. Just because a concept is useless for some roles, doesn't mean it's useless for all roles.

In some cases, the old concept is so defective—maybe an inconsistent concept, an imprecise concept, an immoral concept—that it's better gone. My own view is concepts which are that defective are actually fairly rare, and it's not the the typical case. But in cases where the old concept is not defective, the concept pluralist says, why not keep them both around for different purposes. Then we have a fruitful conceptual apparatus.

The difficulty of conceptual engineering

I was going to say something about the difficulty of conceptual engineering, which Herman stresses in his externalist challenge. My view is that it's clearly possible and clearly difficult. It's a very difficult social project. I think that breaking it into stages helps. The relevant stages here are (i) designing a concept, such as proposing a meaning for a word, and (ii) implementing the concept, say ensuring the word is used to express this concept within a community. The design

project is not that hard. You propose a definition or inferential role or application to paradigm cases. It's not clear that externalism is a major obstacle to that. Implementing it is harder work. First, you have to use it that way yourself—even that's nontrivial. Second, you have to get others to use it that way – that's really hard. And then third, even after it's used that way, you have to ensure that the meaning, the reference, is right.

Externalism I think of as mainly affecting this third step, the step from use to meaning. The externalist will say that narrow use doesn't fix meaning. I think there's various things one can say here. My view is that the central work of the activist project of conceptual engineering can be done even by changing use widely enough. Usually doing that will be enough to change the meaning. If it turns out that it doesn't, changing use, I think, is what we need for many practical purposes. I think this connects to what Jared Riggs will be talking about this afternoon, on different notions of meaning. So I'm not sure that the externalist gap between use and meaning is the biggest obstacle to conceptual activism. A bigger obstacle is the project of changing use, especially the social project of getting others to use a word your way.

Stepping back, I think once you break down conceptual engineering into design, implementation, and evaluation—concept design and concept evaluation are relatively straightforward, or at least don't face the same sort of social obstacles. They suffice for many theoretical purposes. Concept implementation requires changing uses, and is a difficult social project, but it's possible.

Conclusions

Some conclusions. Conceptual engineering includes, or at least should include, de novo conceptual engineering as well as conceptual re-engineering, and heteronymous as well as homonymous conceptual engineering. In my view de novo conceptual engineering is often the most fruitful, especially for theoretical purposes. Homonymous conceptual engineering is also very important, especially for practical purposes that can go well beyond philosophy. I'm strongly inclined think that conceptual engineering doesn't exhaust philosophy. In much of philosophy theses have priority over concepts, and the importance of conceptual engineering in philosophy derives from its role in thesis engineering.

Discussion period

Paul Boghossian: Thanks, Dave. Very useful distinctions. To introduce a new example, to me one of the most prominent examples of de novo engineering is the concept genocide. That word indeed exists. That's a good case where you can ask what de novo means exactly, because the concept has a definition in terms of previously available concepts. So it's compositionally available in some sense. But in some sense highlighted as new, as important for certain purposes. And this connects up with what you were saying about theses. But the idea is, you know, Lemkin noticed that there was a phenomenon that had not been picked out. It had certain features, he thought those features were important for legal purposes, moral purposes, and so on. And so he introduced the concept in order to name that. But it seems as though that—so that general phenomenon, where you notice a phenomenon—of course there are many phenomena—there are murders committed on a Tuesday, you could introduce a word for that—but there, I mean although you might have introduced a new concept, it's not clear what use is the word. So it looks as though this is just, in some sense—I mean science, right? I mean—that what is the distinctive thing has to do specifically with concepts—

Chalmers: I think that conceptual engineering is extremely important in science and in all kinds of social domains. So the question wouldn't even so much be what is so distinctive about concepts here, the question would be what is so distinctive about philosophy here. And then I think the answer might be nothing. I think you know I think conceptual engineering in philosophy is to some considerable extent continuous with conceptual engineering in science, conceptual engineering in social domains, and so on. Nonetheless, all three of those things are extremely interesting for a philosopher to analyze, and the philosophical analysis of conceptual engineering, I think, is a distinctive project for the philosopher.

Boghossian: Yeah, I mean the emphasis in the case of genocide is, as it were, noticing there is a phenomenon.

Chalmers: Phenomenon engineering.

Boghossian: Phenomenon engineering. A phenomenon that has—about which there are important generalizations, and then you need a way of talking about it.

Chalmers: So there's a phenomenon. Well, there's a category, really. You can't just point to the Holocaust and say, there's a phenomenon. It's putting it under a category, a general category, which seems to be the distinctive contribution of a word like 'genocide'. You point to a few paradigm examples and say, look at all the things they have in common. Look at their common moral role. So I guess I think if that's category engineering, that would be the relevant thing.

And you can see categories as things which are out there in the world. Is it property engineering? Well again, we didn't create the property. What happens in this case is we come to stand in a certain epistemological or semantic relation to that property. One way to describe that is to come to possess a concept of that property. So for that reason I think the word 'conceptual engineering' is still apt for this kind of phenomenon.

Paul Horwich: Just following up on this. As Paul said, sometimes you design and engineer the concept by giving a definition. But the skeptical worry is that, if you've got the definition you've already got the concept. So you're not really creating a new thing. What's going on isn't the creation of the concept. The concept's there already. You had it already. It's the implementation problem. You've got a word, and you want people to pay attention, you know, you're directing attention at something by people using the word. So that suggests that the issue is about conceptual activism versus conceptual engineering. And this seems more on the activism side.

Chalmers: Yeah, except I guess I'm inclined to be not so sure that everything you can define from your existing concepts is a concept you have already. Does a random person who happens to have the concepts of modality and the sharing of properties already have the concept of supervenience? Maybe they're in a position to acquire it. I think, you know, Moore and Hare and Davidson and Kim made serious conceptual epistemic progress by articulating that definition, and thereby putting people in a better position to actually possess and deploy that concept.

Horwich: Isn't that the activism side? Getting people to use the concept. I mean, you're right—I shouldn't have said everybody's got the concept.

Chalmers: Okay, there's still the design side. What goes on in the design side—let's say Hare was designing the concept of supervenience. Let's say he never actually got anyone to possess the concept because of externalism. When anyone says supervenience, it means the old meaning of supervenience. I can't remember what the old meaning of supervenience is. I just remember it's used in Quine's autobiography. Anybody remember that? "Necking supervened in the fullness of time as necking will." It was the autobiography of his teenage years. I'm sorry. On this use of supervenience it means something like 'eventuates'. Anyway. Even if after Davidson, Kim, and Moore, supervenience still means that thing, then still Davidson, Kim, Hare, Moore made a contribution by pointing us—by designing that concept, by pointing us to it, to point to all the roles that it can play, even if nobody ever actually comes to possess that concept. I think that would still be some kind of progress. I agree that after that is activism.

Annette Martin: I think I see what you're getting at with this distinction between de novo and re-engineering. But then that immediately sort of raised for me a question about the object

of evaluation. So when we're thinking about deficient representational devices. So in particular, especially thinking about sort of the social context. It strikes me that maybe we should be thinking about frameworks of concepts, as being the deficient thing. So I'm thinking about Miranda Fricker's discussion of the concept of sexual harassment, and how this didn't exist, and thinking, well, part of what was so bad about this wasn't just that there wasn't this concept of sexual harassment, but also the concept that people had of women, and men, and flirting and things like this—this whole cultural framework that's going on. So in introducing the concept of sexual harassment, it seems like that also required some shifting in these other concepts, and so there, again, it seems like maybe the object of evaluation should be more the framework. And then it's like de novo re-engineering. I mean I suppose you can still think of de novo engineering, but it would be a much bigger, radical project.

Chalmers: Yeah. I think you're absolutely right that a lot of the time it is going to be these interconnected frameworks that really get re-engineered. And then you're right—so either you— I mean, well especially if it's going to have to be heteronymous, then you have to introduce a whole lot of new words. One thing that does happen sometimes in some cases is to introduce a common prefix to express all these things that's re-engineered. In teleological terms, we use teleo-this, or things have to be re-engineered in moral terms, you use moral— So there are ways of getting at the project heteronymously, say with something like a prefix, but in those cases I think you're absolutely right that there are particular practical attractions to the homonymous strategy. It's still—I mean, whether it's de novo or re-engineering, I guess the question is, do you think the old framework is fundamentally defective and should be gotten rid of. In some cases I think maybe that will be the case. Maybe in the case of— the whole framework is fundamentally immoral or unjust. Certainly for many purposes we should get rid of it. In other cases I could imagine—you know, here's a re-engineered version of this framework, which is useful, but the old one is still useful for some other purposes, and then you might want to keep both around.

Kate Ritchie: I was wondering about the claim that theses have priority over concepts. I certainly agree that conceptual engineering is not the only thing that philosophy is about. But I was thinking of the difference between projects you might think of as more descriptive versus projects you might think of as more normative. So in the case of supervenience, you might think that defending physicalism requires supervenience. Versus cases where you're saying something like, we need to rethink what race and gender are, or what assault could involve, or like whether injustice could— Katherine Jenkins is pushing a lot that there can be ontological injustice. You could build theses around these, obviously. Sort of like, we ought to use concepts like this because

that'll promote a better world or something. But I was wondering if there might be a difference between the sort of priority theses involve when you have a project that's more descriptive versus more normative.

Chalmers: Yeah. I made this distinction between theoretical and practical projects, and in what I said about the priority of theses I was thinking about the theoretical side. But of course there are theoretical normative projects which sort of mix the theoretical and the practical. Certainly re-engineering the concept of woman or the concept of race has got a very clear social use, one which I don't think really derives from theses. It derives from roles. It derives from the roles that categories of race, womanhood, and so on play in our society. They have major roles, and yeah, we think it would be better in some cases maybe for those roles not to be played at all. That's something like Anthony Appiah's view of race. Or in other cases something different should be to playing those roles. So for the practical uses I don't think I'd want to argue for the primacy of theses. Something more like the primacy of roles.

For the normative theoretical purposes, as you say, we can find theses here. Maybe a normative thesis is that we should use the word 'woman' to express the concept in terms of oppression or the concept in terms of identification. That would be a meta-linguistic normative theoretical thesis. I still think maybe the theoretical importance of the concept derives from the truth of that thesis. At the same time, in normative cases the truth of the thesis goes along with the concept's playing some important practical role. I think there's a way to hang on to the primacy of theses here, but I have to think it through.

X: I'm wondering what you would say to someone like Mark Wilson, whose view seems to be that the whole history of scientific language is simply different kinds of conceptual engineering, and that [unclear] philosophy of scientific language at least is the best place within a traditional semantic theory. Something more pragmatist [unclear]–

Chalmers: I've read a lot of Mark Wilson's work, and it's true that he thinks conceptual engineering is very important, but is it true that he thinks that's somehow– that's all of science, or just that it plays a very important role in science.

X: He doesn't think that there's much for philosophy of language to do except describe the history of science.

Chalmers: I think that's a more general thesis of Mark Wilson's. Not necessarily tied to– in general he thinks that the philosophy of language should be the history of scientific language. All of the interests of language will be brought out by looking at what goes on in actual scientific cases. And then a further thesis–an awful lot of what goes on with scientific language is conceptual

engineering. And then apply the first thesis: the best way to understand conceptual engineering is to look at the history of science. I certainly think that you can get an awful lot of insight into conceptual engineering by looking at history of science. But the stronger thesis, that somehow this is all we should be doing as philosophers, I guess I'm inclined to reject both the first thesis, that there's no more to philosophy of language than looking at history of science—for some pretty obvious and familiar reasons. But I'm also inclined to reject the second thesis. I think there's a lot going on with language in history of science, and conceptual engineering is one central and crucial part of it, but not the only part.

Jackson Kernion: Thanks. Just to pick up on a question that was already asked, I guess I was—in a sense maybe the internalist challenge to conceptual engineering. So one way of seeing all this is that we already have the basic concepts, what we're doing is just we're putting them together in new ways to—well, yeah, I guess the challenge then would be something like, well what we're really doing is we're not conceptual engineering; what we're doing is we're belief engineering. So what we're doing is we are—so the hard part of conceptual engineering is getting people to use the terms. This might just be construed as getting people to have the beliefs we already have. And when you're designing a concept, you might think of that as, well what is guiding me when I'm designing a concept. Well I'm looking at my beliefs, or looking at the world. I'm trying to figure out what category of mine is the fruitful category for generalizations or whatnot. I'm not actually sure this is a challenge, but anyways I just wanted to frame that.

Chalmers: So this is a version of the primacy of theses thesis, except for mental states, where we have the primacy of beliefs thesis. It's like, let's engineer—if we're thinking about engineering mental entities, let's engineer mental entities that can be used in useful beliefs. Is this—are you thinking within philosophy or more generally?

Kernion: I guess—I guess the challenge would be to just like— conceptual engineering as a category distinct from other kinds of belief-changing processes. So you might distinguish between conceptual engineering on the one hand like the arguing over theses— I'm not sure actually that this is correct but I'm curious whether— yeah, whether you think that there is a clear line between, sort of, arguing with theses and conceptual engineering.

Chalmers: I don't think there's a clear line. So I think—I mean one thing we're trying to do in science and philosophy and many other domains is to understand things and figure out the truth. Sometimes it involves recombining old concepts, but sometimes it turns out to get real insight you need a new concept. In the linguistic domain that might involve coming up with a new word; in the mental domain that might involve coming up with a new concept. I guess I'm inclined to

think that—again the importance of the new concept is going to be not just—hey, cool concept, but the work it can do for you mentally. And some considerable part of that will involve beliefs. I mean maybe it will also involve understanding, there may be some domains that might involve connections to action. Maybe it's going to lead to better or more just outcomes. So again I'm inclined to think concepts are going to be of instrumental utility, and one very central part of that will be its role in the formulation of beliefs. And then—I mean you're right that nothing here is absolutely new. It's like, we're trying to figure out the world, sometimes we come up with new categories and new concepts in the in the service of beliefs.

Sigurd Jorem: Yes, you sort of answered this now. I was also puzzled by the primacy of theses over concepts claim you made. And I just wanted to sort of add to the discussion the idea that there's a—for any sort of argument about theses, there's always a standard of goodness or correctness having to do with what concepts you've used to articulate those theses that doesn't have directly to do with the truth of the thesis that you're arguing for. It's like present everywhere, and then it's sort it very hard to—

Chalmers: Could you give an example of this kind of evaluation of concepts that you have in mind?

Jorem: Well, you could criticize a statement for being vague, using a vague concept, and then in that case—I just can't see how you can separate out the concept from the thesis that it's used in.

Chalmers: Yeah, let me try something here. I mean, mean why is the vagueness of the concept used in a thesis a problem here? Because it makes the thesis vague. Just say I offered a thesis which is really precise, and then I did 'and P or not P' on the end, for a vague concept. I don't think you could criticize that for using a vague concept, precisely because it didn't really matter, because the thesis itself wasn't vague—let us assume a semantics where the thesis itself comes out comes out non-vague. There it seems to me it's not such a bad problem. The problem of the—the fault of the vagueness of—the reason why vagueness of concepts matters here is that it brings about vague theses. The reasons why an inconsistent concept might be a problem is they bring about inconsistent on non-existent theses. Insofar as these things don't permeate up to the level of theses, I'm inclined to think those problems with concepts aren't so bad. Now that's probably too strong a claim—I could probably be convinced—maybe there are some concepts that it's just immoral to think with, for example. I mean I'm dubious about whether there are these concepts like 'Boche', which have these ampliative moral inferences built in. But if they did, that might be an example of a concept that it's somehow immoral or problematic to think with. Not because of any roles they play in theses but because of bad things they do to you. Maybe vagueness and imprecision could

be—I mean some people think that, for example, naturalness is an example here, that naturalness is just desirable, in its own right, as something for a concept to express. But I'm inclined to think naturalness is desirable because it's productive and fruitful with respect to theses. It helps you articulate laws and principles. If it wasn't—if we had something natural and something unnatural that were equally productive, then it's not clear to me that unnaturalness would be a problem.

Laura Schroeter: Just a followup on this point. One way in which you could have conceptual effects might be that it triggers certain sort of stereotypes, or you tend to see—even though you've got a determinate extension like with a slur term, you've gotten to see the subject through the lens of a particular stereotype zone. It triggers cognitive effects that are not really truth conditional effects.

Chalmers: Yeah, so that would be sort of a practical effect of certain concepts, given—maybe tied to our nature as non-ideal reasoners. Maybe an ideal reasoner could deploy those concepts without those effects. Maybe you're doubtful about that. I guess I'm inclined to think that the question's going to be, are those effects constitutive of the concept or not? If they're not constitutive, then it's not exactly a defect of the concept so much as the way it's embedded in our own cognitive system. If it is constitutive then maybe we're getting closer to those concepts like 'Boche' and so on.

Herman Cappelen: So I want to hear a little bit more about how you think about the connection between the de novo case and the other. So I was thinking, well, what about that, and to describe roughly how I ended up thinking about it—I think the way you set it up by focusing on this term 'engineering' is a little cheaty, because you know it's really just a label, it's not very descriptive, and you can use other terms, so looking at sort of compositional and how people define the term is—

Chalmers: Externalism, Herman! The words are your masters, not your servants!

Cappelen: So here's a phenomenon, okay? Let me just describe a phenomenon. So there are these philosophers who try to improve concepts like truth or belief or freedom. And that's continuous with things that—the concept of salad has changed over time. In biology the concept of gene, in economics the concept of currency. So these changes happen all the time. That's really interesting: how does that happen, what are the things that happen, what are the things that change, how do those changes happen, to what extent are they intentional, how is topic preserved through those changes. Now I said a whole bunch—you know and the other question is, why do you keep the same lexical item, when is it appropriate to change it. Now I've described a bunch of really hard questions, and I don't think that if I answer those questions, I've answered the de novo

questions. So that's why they're different projects. I could do a really good job answering those questions, and there would still be these unbelievably hard questions about de novo. So they're kind of different.

Chalmers: I totally agree there are distinctive questions about the two cases, and there are also interesting questions about their union that they have in common. I guess I'm interested in all three of those categories. You're especially interested in the questions which are specifically about conceptual re-engineering. That's great. Somebody ought to be interested in it. Carnap's project was conceptual re-engineering.

Cappelen: Do you think the answer to the re-engineering issues I just raised would help you to solve—

Chalmers: I'm almost certain that some of them would help. Give me the list of ten questions.

Cappelen: How do those changes gradually happen in all the examples I gave you, when is topic preserved, when is it appropriate to preserve the lexical item, and when should it be changed. Let's just start with those.

Chalmers: Okay, when is it appropriate to preserve a lexical item and when should it be changed—that's clearly a question which applies to both homonymous and heteronymous conceptual engineering.

Cappelen: But the de novo case doesn't have a lexical item, so the issue doesn't arise there.

Chalmers: Yeah. I think this is going to overlap with the question of to what extent is the old concept useful, and should it be kept around, and to what extent do we end up with plural concepts, and to what extent should we end up with just a single concept and get rid of the other one. And that is very much also a question about when is de novo conceptual—when is pure conceptual re-engineering appropriate, which is getting rid of the old one and replacing it with a new one, versus adding a new concept—when is that appropriate. Anyway, I wouldn't want to deny there are a lot of really interesting distinctive questions about conceptual re-engineering, but I think if you go through, question by question, you'll find (a) de novo conceptual engineering is a really interesting phenomenon; (b) there are a lot of interesting and distinctive questions about that; and (c) an awful lot of questions about conceptual re-engineering and about de novo conceptual engineering actually have a common answer. So it doesn't make sense to study these two things in complete isolation from each other. I guess from your perspective the key issue would be the third one.

[Another reply: it's fine to focus on conceptual re-engineering if you want to, and there are certainly some issues specific to that. But words have power, and “conceptual engineering” is a

powerful phrase. If people had just talked of “concept repair” or “concept re-engineering”, I don’t think the interest would have been so great. The power comes in part because of the suggestion of something that goes beyond mere fixing. I’m trying to vindicate some of that power.]

Robert Rupert: My question requires some setup. I’ll try to be quick about it. I have a question about standards for success in a re-engineering project. So let’s assume that you and Andy were re-engineering the concept of belief in that 90s paper. And then we ask, what good does that do? Clark’s typical sort of response is to say, it opens up new pathways for research in cognitive science, and things like that. But when one goes back and sort of looks at the work of the folks doing the leading work in the 90s on situated cognition and whatnot, it doesn’t look like they really had commitments here. Sometimes they talk about [inaudible] tools, sometimes they talk about it as part of the cognitive system. It doesn’t really much. So I’m worried about the evaluation here, particularly with regard to the scientific case. Do we know enough about, like, how experimental design is generated, or how people get ideas for a quantitative analysis of data, to really make any claims about which of these ways of thinking about belief is what drove the scientific progress. It seems like there might be a sort of disconnect here between—

Chalmers: Yeah, it may be that Andy and I have somewhat different attitudes about this. For Andy the primary purpose was the use in cognitive science. For me that’s one purpose, but among many. And I think what we’re finding in thinking about the extended mind is having—it’s used for all kinds of people dealing with technology and society, for reasons that go way beyond the science. People’s relationship to their smartphones, or their internet search engines. It’s Ned who said that back when we wrote the paper the thesis was false, but it’s since become true. I don’t think that’s because of its role within science. To assess the role within science—I don’t know, I see all this enormous literature these days on 4E cognition, where extended cognition gets at least to be one of the Es. And many people claim that it’s playing a central role in the science. I do agree that many of the points in science where someone appeals to extended cognition could often be reinterpreted as an appeal to embedded cognition. It’s kind of my own view that for any explanatory use of these things, there’s going to be a nearby explanation where it’s done in non-extended terms. It may just be less fruitful. To assess that we really have to engage in the Mark Wilson project of looking at the science. But I do think there are uses outside science.

Gabriel Rabin: Yeah, I wanted to talk about the challenge to de novo conceptual engineering that both Pauls pushed, which was something like, a lot of these paradigms seem to involve a case like rigid designator or supervenience, where you introduce a new word in, but it’s a concept that’s really a concept for the conjunction of some old concepts. And that looks a bit mysterious, so

maybe everything is on the practical side of getting people to use that word. I want to suggest that there's a different type of thing that might go on where you introduce a new concept, and you don't give any type of definition, and you kind of fail at doing so. There might be lots of cases like this in the history of science. There's Newton and Leibniz were wondering about the notion of a limit, and kind of offer these definitions but they aren't very good. Maybe there's a concept that they have, but they can't really articulate it. That seems like there's another type of de novo conceptual engineering, but it doesn't face this kind of challenge, and that's got to be possible. I mean the other alternative is very strong nativism that all our conceptual resources are there from the start, and all we can do is staple them together.

Chalmers: So what's an example of the non-stapling kind of conceptual engineering?

Rabin: It was supposed to be the notion of a limit from Newton and Leibniz. I mean it's much more mysterious how that works. But it's gotta be possible.

Chalmers: Then there was a definition. I mean, there was a concept of a limit and associated things like continuity. And then at a certain point [with Cauchy and Weierstrauss] we get a definition of those things. And I guess it's an interesting question whether it's a new concept. I think probably to some extent it's at the very least a precisification. Now the new concept, the epsilon-delta definition of limits and continuity—well, I mean it's not so clear to me that's different in kind from supervenience. It's something that someone has stapled together from things we understood. You know, the math of epsilons and deltas is pretty well understood. What's new is (a) the stapling and (b) the connecting them to that role, to that old, fuzzy notion we had of limits and continuity. I think that's kind of analogous to what goes on with, say, this precise concept of supervenience, and at least attempting to connect that to the old, fuzzy concept of materialism or physicalism. So there's something which is engineered, which is fairly precisely. What's kind of fuzzy and vague is the role, and what's fruitful is somehow the connection between the two.

[To clarify, I'm not saying that de novo conceptual engineering always involves definitions composed from old concepts. Often it involves something fuzzier, like pointing to paradigm cases and gesturing toward something in common, or formulating a big bold theory using a new concept, or just gesturing at a role we'd like a concept to play. This isn't just stapling old concepts. At the same time, it's an interesting question whether these newly engineered concepts are ever deeply and fundamentally new. I'm at least tempted to think that they are usually grounded in old concepts in some non-stapling way. Maybe they're pinned down by inferential connections to old concepts. I've talked about that sort of thing here in other work. But the idea of de novo conceptual engineering is meant to be neutral on this issue. New concepts engineered by stapling old ones

count as de novo, new concepts grounded in inferential connections to old concepts count, and so do new concepts not grounded in this way, if there are any.]