Synthese (2021) 198 (Suppl 7):S1507–S1527
https://doi.org/10.1007/s11229-019-02096-x

S.I. : KNOWLEDGE AND JUSTIFICATION, NEW PERSPECTIVES

Conceptual engineering, truth, and efficacy

Jennifer Nado[1]

Received: 18 October 2017 / Accepted: 16 January 2019 / Published online: 23 January 2019

Abstract
Traditional views on philosophical methodology characterize our primary philosophical goal as production of a successful conceptual analysis. The notion of conceptual analysis, however, faces several challenges—from experimental philosophy to more traditional worries such as the paradox of analysis. This paper explores an alternate approach, commonly called conceptual engineering, which aims at recommending conceptual revisions. An important question for the conceptual engineer is as follows: what counts as a case of successful conceptual engineering? What sorts of revisions are permitted, and what sorts are too revisionary? In this paper I examine 'functional' approaches to conceptual engineering, ultimately arguing for a 'radical' functionalism according to which even revisions which 'change the subject' are permitted, and successful re-engineering is constrained only by the requirement that continuity in needed functions of a pre-engineering concept be maintained somewhere in the postengineering conceptual scheme. I further argue that this approach suggests a heightened role, in metaphilosophical discourse, for a neglected epistemic goal—conceptual efficacy.

Keywords Conceptual engineering · Conceptual analysis · Intuition · Explication

1 The Standard Model of philosophical methodology

The following is a well-worn, familiar description of what philosophers do. First, philosopher McA selects a philosophically interesting concept C ('freedom', 'consciousness', 'good', etc.), and proposes a theory which purports to delineate the conditions under which something counts as C. Ideally, McA's theory will take the form of a biconditional, the left-hand side of which contains C, and the right-hand side of which contains necessary and sufficient conditions for being C. Thus, McA's theory

✉ Jennifer Nado
jennifernado@gmail.com

[1] Department of Philosophy, University of Hong Kong, Pokfulam, Hong Kong

(if successful) should enable us to determine, for any hypothetical case, whether it falls under C or does not.

Next, philosopher McB challenges McA's theory by producing an imagined case (sometimes quite bizarre or complex). This should be a case which McA's theory deems to be C, but which intuition deems to be not C (or vice versa). If McB can generate such a case, this counts strongly against McA's theory—perhaps strongly enough to warrant its rejection. As the history of twentieth-century philosophy shows, even a single such intuitive counterexample may condemn an otherwise elegant theory, sending McA back to the drawing board.

The method by which McB challenges McA's theory is often called the 'method of cases'; McA's theory itself is standardly called an 'analysis'. An analysis of what, you might ask? Well, the traditional answer would be: an analysis of the concept C. In other words, a conceptual analysis. Philosophers, so the story goes, are in the business of producing and testing conceptual analyses. And intuition, generally via the method of cases, is our primary source of evidence for the success (or more commonly, the failure) of an analysis. We might call this general picture of the goings on of philosophy the 'Standard Model' of philosophical inquiry.

The Standard Model's image of philosophy's goals and basic methods has been questioned in various ways, particularly in recent years. The most headline-grabbing worries stem from critiques of intuition—there has been, over the past 25 years or so, a growing sense of doubt about whether intuitions are reliable enough to support a project like conceptual analysis. This doubt antedates, but has been spurred on in large part by, the emergence of the 'negative' strand of experimental philosophy. Over the past two decades, negative x-phi has generated substantial empirical evidence that our intuitions are pushed about by such irrelevant factors as cultural background, order effects, personality, and more.[1] If these doubts about intuition's epistemic credentials are warranted, then the value of the methods portrayed by the Standard Model is called into serious question.[2]

Even for those who do not share such doubts about the reliability of intuition, there are numerous reasons to suspect that the Standard Model is more than a little problematic—from worries about analytic truths, to puzzlement surrounding the paradox of analysis, to doubts about whether 'classical' necessary-and-sufficient-definition models accurately capture the structure of concepts.[3] And, of course, perhaps the most damning criticism of the Standard Model: that its methods seem to have failed, over more than two millennia, to generate even a single successful instance of analysis. Even 'x is a bachelor iff x is an unmarried man' seems to have potential counterexamples in male infants, long-term cohabitation, and the Pope. If the aims of philosophy

[1] For representative examples, see e.g., Weinberg et al. (2001), Machery et al. (2004), Swain et al. (2008), Feltz and Cokely (2008), Tobia et al. (2013) and Buckwalter and Stich (2014).

[2] One reaction here is to deny that the Standard Model accurately captures actual philosophical practice. See for instance Cappelen (2012) and Deutsch (2009, 2010, 2015).

[3] For a classic (though older) summary of the philosophical and psychological work on this last issue, see Laurence and Margolis (1999).

are those suggested by the Standard Model, then we have made depressingly little progress towards achieving them.[4]

But there is an alternative. A growing number of philosophers have begun to suggest that we depart from the Standard Model and its dogged pursuit of conceptual analysis, towards a more revisionary venture—conceptual *engineering*. Conceptual engineers aim to improve or to replace rather than to analyse; to create rather than to discover. While conceptual analysts are interested in the concepts we *do* have, conceptual engineers are interested in the concepts we *ought* to have. Their project is prescriptive rather than descriptive.

Revisionary projects are, of course, nothing new in philosophy, and in-depth explorations of what is now called 'conceptual engineering' date back at least as far as Carnap's proposed methodology of 'explication' (Carnap 1950). But we seem to be on the cusp of entering a conceptual engineering renaissance: there has, within the past few years alone, been an exponential upswing in publications discussing Carnapian explication, 'conceptual ethics',[5] and various other brands of revisionism.[6]

Revisionary projects have a tendency to emerge wherever there is a sense that our current concepts are somehow non-optimal. Many of our concepts are, for instance, imprecise or vague—this perceived flaw motivated not only Carnap's method of explication, but also other instances of 'ideal language' approaches found among early practitioners of analytic philosophy. Many of our current concepts appear unsuited to certain technical contexts—prompting, for instance, the introduction of specialized uses of English terms in fields like formal epistemology ('credence', e.g.), as well as the departures from natural language needed to generate truth-functional logical operators (as in the material conditional). A concept might be defective because empty—thus prompting eliminativist views, as in philosophy of mind or metaphysics. A concept might be defective because incoherent, or paradox-producing—such as, arguably, the naïve notion of truth. A concept might even be defective because politically or socially non-optimal, as suggested e.g., by Sally Haslanger in her revisionary definition of 'woman' (Haslanger 2000). Undoubtedly, the recent wave of anti-intuition sentiment is leading many revision-friendly philosophers to suspect even deeper and more wide-spread deficiencies in our conceptual repertoire.

I'll use 'conceptual engineering' as a very broad umbrella term that encompasses all of the instances of revision just mentioned. 'Conceptual engineering', as I use the term, covers any project which aims to in some sense or another repair defects in our

[4] Of course, all of these worries for the Standard Model have had their fair share of discussion, and potential responses to each are available. There has even been a 'new wave' of defenses of conceptual analysis, under the banner of the so-called 'Canberra plan' (see especially Jackson 1998). Yet the sheer volume of concerns with the Standard Model ought still to give us pause—we ought, I think, to consider whether alternative approaches to the aims and methods of philosophy might skirt the quagmire against which proponents of the Standard Model struggle.

[5] This terminology was coined by Burgess and Plunkett (2013a); it is at least roughly synonymous with 'conceptual engineering', being concerned with prescriptive questions surrounding concept selection. Though see Cappelen and Plunkett (forthcoming) for a more detailed look at the terminology.

[6] Several of these will be discussed in later sections; for only a few of the many which have been neglected due to space constraints, see e.g., Maher (2007), Kitcher (2008), Olsson (2015), Eklund (2015), Plunkett (2015), Pinder (2017a) and Machery (2017).

conceptual system.[7] The primary focus of the current paper will be a question that looms large for any aspiring conceptual engineer: what counts as a *successful* instance of engineering? I'll be exploring the prospects for a 'functionalist' approach to this question. Roughly, a functionalist takes success in conceptual engineering to be a matter of how well the engineered concept fulfils certain functions, aims, or purposes. I'll explore a few extant versions of this approach, ultimately arguing for a 'radical' functionalist view which permits more drastic revisions than most other accounts on offer.

I'll then go on to discuss what I find to be an interesting epistemological implication of the more 'radical' functionalist approach I favour—and potentially of other versions of conceptual engineering, as well. In short, I'll claim that conceptual engineering invites a reorientation of our philosophical aims away from those the Standard Model tends to presuppose—away, that is, from standard epistemological 'goods' such as truth and knowledge. A conceptual engineer, I claim, should instead view success in philosophy primarily in terms of conceptual *efficacy*. Conceptual efficacy is the chief aim of philosophy; truth and knowledge, by contrast, are secondary goals at best.

2 Success in conceptual engineering: the semantic approach

At least in rough outline, we know what it means to get a conceptual analysis 'right'. It requires that the proposed analysis capture the meaning or content of the analysed concept: that all cases, real or imaginary, that fall under the analysans also fall under the analysandum, and vice versa. There must be, in other words, no counterexamples. But conceptual engineering is not subject to the same requirement. 'Counterexamples' are part of the package; the whole point of the enterprise is for the revised concept to depart from its predecessor. This is, indeed, why conceptual engineering is such an attractive alternative to the Standard Model—since it is permitted to depart from the classifications offered up by intuition, it is not subverted by the apparent unreliability of intuition. Engineering is a means to *mitigate* the unreliability that threatens the Standard Model—the points where intuition is thought to be in error are exactly those in need of the engineer's metaphorical hammer or wrench.

But if we cannot measure the success of a case of conceptual engineering by counting counterexamples, how are we to measure it? Clearly, it will have something to do with removing the conceptual flaws that prompt us to engineer in the first place—reducing vagueness, for instance, or modifying the concept for use in a technical context. But this is arguably not the only requirement. Engineers and analysts both overwhelmingly agree that there must be limits to revision; revisions cannot be *too* revisionary. On the analyst side, this conservativism has led some to doubt the prospects for any satisfactory measure of success in conceptual engineering.

The worry here was most famously articulated by P. F. Strawson in an exchange with Carnap over the merits of explication. Explication, in brief, is Carnap's sug-

[7] There's a lot of unpacking needed here, particularly with regard to the role of 'concept' in conceptual engineering. Cappelen (2018), for instance, objects to the notion that conceptual engineering trucks in concepts at all. I'm less bothered by the 'conceptual' portion of the label, myself, but nothing in the current paper will hang on untangling this particular knot.

gested method for replacing inexact, pre-theoretic terms with more exact versions, thus improving their suitability for use in e.g., logic and the sciences. Strawson, however, argued that to explicate our pre-theoretic terms in the formal, precise way that Carnap intended "is to do something utterly irrelevant—is a sheer misunderstanding, like offering a text-book on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart" (Strawson 1963, p. 504). Explication, in other words, is accused of missing the point—of changing the subject, and thereby failing to answer the philosophical questions that prompted our inquiry in the first place.

Though Strawson was discussing explication, essentially the same worry can be (and has been) raised for engineering projects generally. Conceptual engineering seems to run the risk of being *too* revisionary, and thereby losing sight of the concepts we started with. Thus (for instance), when Haslanger proposes a revisionary definition of 'woman', she is simply no longer talking about women. One can take this concern quite far, especially if one holds that a concept's meaning, intension, or what have you is essential to it. On such views, 'revising' a concept turns out to be impossible. Any change in meaning results in a different concept.

One finds this 'Continuity Problem'[8] cropping up repeatedly in discussions of revisionary projects, even outside of the conceptual engineering literature proper; I'll provide here just a few examples. The first, from Jackson, who has voiced the worry in several places:

> [I]f we give up too many of the properties common sense associates with belief as represented by the folk theory of belief, we do indeed change the subject, and are no longer talking about belief. The role of the intuitions about possible cases so distinctive of conceptual analysis is precisely to make explicit our implicit folk theory and, in particular, to make explicit which properties are really central to some state's being correctly described as a belief. For surely it is possible to change the subject, and how else could one do it other than by abandoning what is most central to defining one's subject? (Jackson 1998, 38).

Another, from Alvin Goldman:

> Whatever else epistemology might proceed to do, it should at least have its roots in the concepts and practices of the folk. If these roots are utterly rejected and abandoned, by what rights would the new discipline call itself "epistemology" at all? It may well be desirable to reform or transcend our epistemic folkways… [b]ut it is essential to preserve continuity; and continuity can only be recognized if we have a satisfactory characterization of our epistemic folkways (Goldman 1993, p. 272).

And finally, even experimentalist critics of intuition have voiced the same concern:

> Philosophical practice is not concerned with understanding the nature of knowledge (or belief, freedom, moral responsibility, etc.) in some technical sense, but

[8] Michael Prinzing (2018) uses the term 'Discontinuity objection' for the same worry; it's also often simply called 'Strawson's objection', 'Strawson's challenge', or something similar.

> of knowledge as the concept is ordinarily understood outside of strictly philosophical discourse and practice. If it were concerned only with the technical sense of the concept, it would be divorced from the concerns that led us to philosophical investigation of the concept in the first place and its verdicts would have little bearing on those initial concerns (Alexander and Weinberg 2007, p. 58).

Examples could surely be multiplied, but I trust the general gist is clear. There is a common supposition that conceptual engineers must preserve continuity with the pre-engineering concept, else they fail to address our original philosophical questions. Some revision of the pre-engineering concept is allowable in the course of a philosophical investigation, but only some.

Some of the more radical among us (myself included) may be unfazed by the idea that conceptual engineering might 'change the subject'. If our initial concepts are problematic, then perhaps our initial philosophical questions aim at the wrong targets—like an old chemist puzzling over why the release of phlogiston occasionally prompts a gain in mass in a burned object. But even those who are comfortable with a rethinking of our philosophical problemata should still give serious weight to the Continuity Problem. For, as Jackson has pointed out, the problem is not *just* that revision threatens to change the subject—it is that revision threatens to trivialize philosophical problem-solving.

> If I say that what I mean – never mind what others mean – by 'belief' is any information-carrying state that causes subjects to utter sentences like 'I believe that snow is white', the existence of beliefs so conceived will be safe from the eliminativists' arguments. But [I will not] have much of an audience. I have turned interesting philosophical debates into easy exercises in deductions from stipulative definitions together with accepted facts (Jackson 1998, p. 31).

There is something to this worry, even for fans of radical revision. We can't, say, resolve debates over free will by engineering just any old concept and calling it 'free will', as follows:

> Free will: x possesses free will iff x is an H2O molecule.

Not even the staunchest revisionist would take the above proposal seriously. But consider for a moment—why not?

In a sense, there is nothing wrong with the proposed concept. It is not defective in any of the ways revisionists have been concerned with: it is not vague, nor is it empty, nor is it incoherent. It tracks a natural kind. What more could we want from a concept? But nonetheless, it clearly does not succeed as an engineered replacement for the concept of free will—presumably, because it lacks continuity with the original. Even if we are happy to in some sense 'change the subject', we need an account of successful engineering that rules out trivializing proposals like the one just given.

It looks, then, as though a satisfactory account of engineering success will involve multiple desiderata—one (or more) to measure the degree to which the flaws of the original concept have been ameliorated, and one (or more) to measure the continuity that has nonetheless been maintained. This is, in fact, exactly how Carnap structured his own account of successful explication. For Carnap, a successful explication will

possess the following characteristics: *similarity* to the explicandum, *exactness*, *fruitfulness* (in the sense of enabling the formulation of universal statements such as scientific laws or logical theorems), and *simplicity*. It is the similarity desideratum, of course, that ensures continuity. Here is how Carnap states the criterion:

> The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted (Carnap 1950, p. 7).

Carnap notes, for instance, that we are permitted to alter the extension of a concept in order to achieve a gain in one of the other desiderata. Thus, the scientific concept of 'fish' (which Carnap refers to as 'Piscis') improves on the fruitfulness of the prescientific concept by the removal of aquatic mammals such as whales.

The above example suggests that the way to read Carnap's similarity desideratum is in terms of *meaning*. In other words, the required continuity is to be maintained by ensuring that the meaning of the explicandum is similar enough to the meaning of the explicatum. Though Strawson clearly wasn't satisfied with the continuity afforded by the desideratum as Carnap stated it, we might read this as a disagreement over degree rather than kind—over how much similarity is required—with Strawson holding doubts that substantial gains in exactness or fruitfulness can be made without falling below the required meaning similarity threshhold.

This take on the Strawson/Carnap debate illustrates what we might call the 'semantic' approach to maintaining continuity. The semantic approach holds that maintaining continuity requires preserving similarity in the meaning or content of the target term or concept. Of course, there are more details to be worked out—one might imagine debate over, e.g., whether the required meaning similarity involves extension, intension, or what have you. But regardless of the particulars, the semantic solution appears to have a nice corollary: it suggests that any successful instance of conceptual engineering will need to be preceded by a fair amount of conceptual analysis.[9] After all, we plausibly need to know a good deal about the meaning of the pre-engineering term/concept in order to ensure that we are maintaining the needed similarity.

Perhaps it is not surprising, then, that many fans of the Standard Model seem to have assumed that revisionary projects must solve the Continuity Problem via some form of meaning similarity. This seems to be what Jackson and Goldman have in mind, for instance. On the other side of the aisle, many fans of Carnapian explication or other forms of conceptual engineering also appear to take a semantic approach to continuity. Some have pointed out that even Carnap himself suggested that some amount of clarification of meaning (of the explicandum) is a prerequisite for successful explication (Justus 2012; Shepherd and Justus 2015; Novaes and Reck 2017). Going further down this route, Schupbach (2015) even suggests that we might pursue what he terms "Oppenheimian explication", which is a variant of Carnap's project (inspired by Kemeny and Oppenheim 1952) which retains the same desiderata but prioritizes maintaining similarity (of meaning) over producing gains in fruitfulness. Cappelen (2018), meanwhile, takes a somewhat more flexible semantic approach by claiming

[9] Though see Schupbach (2015) for an argument that experimental philosophy can fill this role.

that we can maintain similarity via maintenance of *topic*. This topic-based approach appeals to the general linguistic phenomenon whereby two people can be talking about the 'same thing', or indeed 'same-saying', even when their terms vary somewhat in intension or extension.

If we embrace a reasonably strong semantic restriction on revision, we may be led to a view of conceptual engineering that is, while revisionary, also fairly limited. We may come to think that the goal of conceptual engineering is just to do a touch of 'clean-up'—to precisify the somewhat messy concepts handed to us by folk usage, possibly making a few extensional changes along any jagged boundaries, while carefully leaving their core meanings intact. Most versions of the semantic approach would suggest that if there is only a small degree in overlap in extension between the pre- and post- engineering concepts, we have failed. Or, that if certain core cases do not fall under the post-engineering concept, we have failed. Conceptual engineering thereby begins to look quite beholden to analysis; the resultant picture of philosophy begins to look not terrifically different from the Standard Model.

To some that will seem a favourable aspect of the view. But there are reasons why a more adventurous engineer might prefer to leave open the possibility of more radical interventions. First, insofar as we conceive of conceptual engineering as a matter of improving our conceptual system as a whole, the invention of novel concepts counts as engineering. In philosophy, this would include e.g., the introduction of terms like 'supervenience' or 'haecceity'. No issue of continuity even arises for such cases. But even leaving such cases aside, it's not obvious that conceptual engineering should be limited to mere revision. In some cases, indeed perhaps many, a concept may be defective enough to warrant outright replacement.

Take an eliminative materialist position on propositional attitudes, for instance. On such proposals, mental categories like belief and desire do not exist, and the corresponding concepts and terms should be replaced—likely by some form of neuroscientific category. Eliminativists don't, and shouldn't, hold that the replacement categories must e.g., retain core platitudes of folk psychology, or preserve sufficient overlap in extension with folk psychology's ontology. For the eliminativist, the platitudes are false and the extensions are empty. Changing the subject is the goal. Of course, the eliminativist arguments for the falsity of folk psychology or the emptiness of its terms may be incorrect, but that's beside the point. What matters is that the strategy itself—replacing radically defective concepts with ones that 'change the subject'—should be a permissible strategy for an engineer. It would be perverse to agree with the eliminativist's claims about the inexistence of belief and desire and then go on to fault her for failing to preserve similarity of meaning in her successor concepts.

And yet, something must be preserved. If an eliminativist suggested rejecting all folk psychological mental categories and replacing them with concepts expressing, say, connectives of propositional logic, we'd rightly object. So what is the required element of continuity? Perhaps it is something semantic; perhaps it is some feature of meaning not captured by extension or by core platitudes. But must it be?[10]

[10] I would also point the reader to a fantastic argument in Brun (2016) as to why similarity cannot plausibly be cashed out extensionally.

3 Success in conceptual engineering: the functional approach

Discussions of Carnapian explication, conceptual ethics, and conceptual engineering are rich with allusions to the *functions*, *purposes*, or *aims* of our concepts. In some cases such terminology is used more or less in passing; in other cases the relevant engineers are explicit in endorsing the role of function-fulfilment in successful engineering. I'm of the opinion that some form of functionalist approach to conceptual engineering is the correct way of dealing with the continuity problem. In other words, I hold that sufficient continuity is provided by continuity of function. *Very* roughly, a successor concept is 'similar enough' to the original concept so long as it serves the same function(s). This approach neatly ties the continuity desideratum to the improvement desideratum: to improve on a concept is to produce a concept which better serves a given function of the original, where that very shared function provides the needed continuity between pre- and post- engineering concepts.

In this section, I'll examine some possibilities for fleshing out the functionalist view. One central issue here will be the question of what exactly the function of a concept *is*. A second, closely related issue will be the relation between a concept's function and its meaning; that is to say, the question of whether a functional approach is a variety of semantic approach. As a preview, I'll argue in favour of a radical functionalist approach which eschews any semantic desideratum on continuity.

Before we get to all that, though, it's worth noting how deeply the functionalist view is embedded in the revisionist tradition. Perhaps the most explicit endorsement of functionalism prior to the current engineering renaissance is due to Quine, in *Word and Object*:

> We do not expose hidden meanings, as the words 'analysis' and 'explication' would suggest; we supply lacks. We fix on the particular *functions* of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms to our liking, that fills those *functions*. Beyond those conditions of partial agreement, dictated by our *interests and purposes*, any traits of the explicans come under the heading of 'don't-cares' (Quine 1960, p. 238, emphasis mine).

And later:

> [E]xplication is elimination. We have, to begin with, an expression or form of expression that is somehow troublesome. It behaves partly like a term but not enough so, or it is vague in ways that bother us, or it puts kinks in a theory or encourages one or another confusion. But also it *serves certain purposes* that are not to be abandoned. Then we find a way of *accomplishing those same purposes through other channels*, using other and less troublesome forms of expression. The old perplexities are resolved (Quine 1960, p. 239, emphasis mine).

Carnap also seems to have functional continuity in mind in his response to Strawson:

> A natural language is like a crude, primitive pocketknife, very useful for *a hundred different purposes*. But *for certain specific purposes, special tools are more efficient*… If we find that the pocket knife is too crude for a given purpose and

> creates defective products, we shall try to discover the cause of the failure, and then either use the knife more skillfully, or *replace it for this special purpose by a more suitable tool, or even invent a new one*. [Strawson's] thesis is like saying that by using a special tool we evade the problem of the correct use of the cruder tool (Carnap 1963, p. 938, emphasis mine).

In both cases, there is a suggestion that we may go beyond revision, to outright replacement—accompanied by no obvious concern about 'changing the subject'.

Even Strawson himself seems to suggest that what really lies behind the 'changing the subject' worry is a concern with our concepts' ability to fulfil certain important functions. Strawson writes:

> Language has many other *employments*. We *use it* in pleading in the law courts; in appraising people's characters and actions; in criticising works of art; in recounting our states of mind; in getting people to fetch things; in narrating histories; in describing what things look and sound and feel like; in entering into engagements with one another; in identifying people – and so on… it seems in general evident that the concepts used in non-scientific kinds of discourse could not literally be replaced by scientific concepts serving just the *same purposes* (Strawson 1963, p. 505, emphasis mine).

The real worry behind Strawson's problem, in other words, seems to be that in the process of improving a concept's suitability to fulfil the purposes of the sciences, we thereby decrease its suitability to fulfil the purposes of ordinary discourse.[11]

Turning to more contemporary work, we encounter a number of apparently functionalist takes on engineering. They do, however, vary somewhat in their characterizations of both the project being pursued and the nature of the functions or purposes involved. I'll here look at a few examples which I find to be both well-developed and illustrative of the various options one might take in attempting to articulate a functionalist position.

At one end, we find a functional approach that I hesitate to label 'engineering' at all: that of Edward Craig in his *Knowledge and the State of Nature* (1990). Craig is there engaged in what he calls "conceptual synthesis", or a "practical explication of knowledge". The aim is to "take some prima facie plausible hypothesis about what the concept of knowledge does for us, what its role in our life might be, and then ask what a concept having that role would be like, what conditions would govern its application" (Craig 1990, p. 2). Craig's suggestion is that the primary purpose of the knowledge-concept is to help us identify good informants.

But despite his use of the term 'explication', Craig's project is not obviously revisionary:

> Carnap's intentions were normative, the establishment of the concepts fit to form the rational basis of the unified science, whereas mine are the more purely theoretical ones of shedding light on the nature and origins of present practice (Craig 1990, p. 8).

[11] For another good alternative interpretation of Strawson's objection, see Pinder (2017b).

Craig's project is not traditional analysis, but the foregoing quote suggests that the goal is primarily descriptive—it is something like an ethnography of our knowledge-concept. If this is correct, then Craig's project, though of interest in its own lights, does not fall under conceptual engineering as I have defined it.

Turning to more clearly revisionary approaches, one possibility for clarifying the notion of a concept's function is to tie functions to meanings—perhaps by appeal to teleosemantics or inferential role semantics. Fisher (2015), for instance, argues in favor of what he calls "pragmatic conceptual analysis", which "seeks an explication that will best preserve the patterns of beneficial usage for a given concept" (Fisher 2015, p. 414). This project is then motivated by its fit with a teleosemantic view on meaning. As with Craig, Fisher's view is in a sense not a revisionist view at all; since Fisher holds a concept's meaning to be determined by those applications that produce said beneficial usage, pragmatic conceptual analysis is "mere clarification rather than revisionary stipulation" (Fisher 2015, p. 416). Despite this sentiment, Fisher writes primarily as a revisionist, so it's worth treating his proposal as such. Taken in a revisionist spirit, Fisher's view would yield a solution to the continuity problem that is both functionalist and semantic. Since a successful instance of engineering preserves function, and since function is determinative of meaning, successful engineering thereby preserves similarity of meaning.

Though most authors are not quite as explicit in endorsing a particular metasemantic view that ties functions to meanings, there are other accounts that suggest some form of semantic link. For example, Prinzing (2018) argues that the proper solution to the continuity problem is preservation of function, and takes preservation of function to be preservation of the 'essential features' of a concept. Prinzing thus takes concepts to be individuated by their function, and claims that revisions that preserve function (and thus concept identity) do not 'change the subject'. He notes the similarity between his own view and inferential role semantics, suggesting that appeal to function might well provide a means for identifying those 'privileged' parts of a concept's inferential role that are essential rather than accidental. On at least one possible reading, then, Prinzing's account is another case of a semantic functional view—or, at the very least, of a functional view that requires preservation of concept identity, thereby avoiding change of subject.

However, a functionalist that holds that functional similarity is needed to avoid 'changing the subject' (whether this is construed semantically or not) risks ruling out a number of *prima facie* legitimate revisionary projects. First, as noted in the previous section, in many cases radical revisions which quite clearly change the subject are not only permissible, but salutary. Should eliminativists be right about the failings of folk psychology, the proper response would be to change the subject, not to retain-but-revise the concepts of 'belief' and 'desire'. The case of obsolete scientific concepts, such as phlogiston, is another case in point. The move from 'phlogiston' to 'oxygen' changed the subject. Prinzing notes this case, and categorizes it as a case where the Continuity Problem (or in his terms, the Discontinuity Objection) does not apply—as a case where it 'is no objection at all' (Prinzing 2018, p. 856). I think, however, that this division between cases where identity-preservation is the means for preserving continuity and cases where we waive the requirement simply isn't needed, or justified. Continuity is still required in cases where the subject has changed. Had a bold chemist

proposed replacing 'phlogiston' with 'carburetor', things would not have gone quite so well. Why not just claim that continuity of function can be maintained without identity?

Accounts of conceptual engineering which forbid changing the subject also threaten to limit our engineering ambitions in other, less obvious ways. Suppose in the process of engineering we split one concept into two—perhaps following Scharp (2013) in substituting two notions of truth with different inferential properties (dubbed 'ascending truth' and 'descending truth') for our paradox-producing pre-theoretic notion. Have we changed the subject? Well, ascending truth and descending truth are not the same concept. Nor do I 'say the same thing' when I label a sentence 'ascending-true' as I do when I label it 'descending-true'. So a puzzle arises here rather like that for cases of fission in personal identity—when A divides into B and C, and B and C are not identical, what do we say of the relationship between progenitor and descendants? Is this proposed fission within the bounds of legitimate revision?[12]

I see no good reason why it shouldn't be, regardless of the semantic or identity relations that hold post-revision. Yet if I 'change the subject' when I replace an 'ascending-true' statement with a 'descending-true' one, can we make sense of the idea that neither 'ascending-true' nor 'descending-true' change the subject away from truth? If a prohibition on changing the subject is a crucial part of engineering, there will be a host of similar cases to puzzle over. There are cases of fusion—what do we say when an engineer proposes combining two pre-theoretic concepts into a single successor? There are cases of partial fission, where the pre-engineering concept is retained for most uses and a successor concept is introduced to replace only certain specialized uses; many technical terms plausibly fall under this category. And so on.

I have been here largely assuming that semantic accounts, functional or otherwise, aim to retain conceptual identity (or at least, following Cappelen, identity of some related semantic notion like 'topic'), thereby providing a similarity threshold that ensures that revisions do not change the subject. This is how I think most such accounts are naturally read, but one could in principle develop a semantic continuity criterion, perhaps couched in terms of function, which permitted outright, subject-changing replacement.[13] Perhaps descending truth is not the same concept as truth simpliciter, but one might argue that the shared element of inferential role nonetheless provides sufficient similarity of meaning. But there are even more radical cases of *prima facie* acceptable engineering that will be difficult to square with even such a permissive semantic approach.

Consider cases where engineering involves a move from one entire *set* of concepts to a second set, as in the replacement of one classificatory system with another. In such cases, there might well be elements of the pre-engineering set which have no

[12] Prinzing notes this case, and claims that the result is a replacement rather than a revision. I suspect he would thus group it with 'phlogiston' as a case where the continuity problem is waived. As with the phlogiston case, I think the proper conclusion is that identity preservation is a red herring.

[13] An interesting approach which might fit into this category is that of Brigandt (2010), who argues for recognizing a concept's 'epistemic goal'—that is, "the kinds of inferences and explanations that the concept is intended to support" (Brigandt 2010, p. 24)—as an aspect of its content alongside reference and inferential role. He goes on to use this notion to give an account of rational semantic change, allowing that reference or inferential role can change so long as this is warranted by the concept's epistemic goal.

clear individual successor in the post-engineering set. An example here might be the move among many biologists from a traditional Linnaean-style taxonomy to one based on clades; while the former system reflects morphological similarities between species, the latter classifies by reference to shared common ancestry. On a cladistic taxonomic system, there is for instance no clear successor to 'fish'—the smallest clade that includes all fish (Euteleostomi) also includes all tetrapods, and thus includes humans. We might hold Carnap's Piscis to be within the bounds of permissible revision of extension, but Euteleostomi doesn't stand a chance.[14]

If we simply speak of functions, with no attempt to tie a concept's function to its meaning, many-to-many revisions/replacements aren't obviously inherently problematic—a set of concepts might have a function or set of functions, all of which are preserved in the set of concepts which serves as replacement. It is less comfortable to retain talk of meaning continuity in such cases. A non-semantic notion of function also faces no particular difficulty in characterizing instances of engineering that are one-to-many, or many-to-one, or which serve as only partial replacements. So why all this fuss about preserving meaning?

There is one final reason why a functional conceptual engineer ought to resist tying functions to preservation of meaning or concept identity. It's plausible to view conceptual engineering as not merely a matter of improving on a given function, but in questioning and critiquing the functions of our concepts, and potentially altering or abandoning those functions. This is a revisionary aim emphasized by, for instance, Burgess and Plunkett (2013b), Plunkett and Sundell (2013), Haslanger (2000), and Thomasson (2017). Yet if functions are too closely tied to meanings, and if change of meaning is prohibited, this sort of meta-revision will be rather arbitrarily limited.

An example from Haslanger's work may illustrate the point. Haslanger's revisionary definition of 'woman' is primarily aimed at identifying and remediating oppression; her concept 'woman' is (in short) that of a person who faces subordination on the basis of perceived biological features indicating a female role in reproduction. Identifying such oppressed individuals is, presumably, not an identity-constitutive function of our pre-theoretic concept 'woman'. Haslanger's concept jettisons certain plausibly much more central functions of our pre-theoretic concept which she finds to be either scientifically indefensible or socially problematic, such as delineating a (supposed) natural biological category, or reinforcing certain traditional social roles. Has Haslanger changed the subject? Probably. But she has done so in a principled fashion, rejecting functions she takes to need rejecting while retaining uses that still hold value.[15]

[14] A semantic engineer might argue that semantic similarity desiderata only apply when we aim to retain a term post-revision, thus allowing that Euteleostomi is legitimate because the category does not retain the label 'fish' (some sections of Cappelen (2018) suggest such a strategy). Nonetheless, the examples in this section are instances of conceptual repair, not novel introductions like 'haecceity'. As noted earlier, continuity is required for such projects. The semantic engineer doesn't have an obvious account of the needed continuity; the functionalist does.

[15] It's worth noting that functions might be rejected for reasons other than social justice, too. For instance, the obsolete measurement 'league' originally referenced the distance a person could walk in an hour. Thus, one plausible function of the 'league' concept was to quickly provide a rough estimate of walking time for long distances. However, in an era of cars and planes, where few of us walk more than an hour at a stretch on any regular basis, it's quite arguable that we no longer need a long-distance measurement calibrated to

4 Radical functionalism and the nature of functions

To sum up: I hold that functional continuity demarcates the limits of permissible revision, but I see no reason to tie functional continuity to meaning or to concept identity. An engineer can 'change the subject', replacing the pre-engineering concept with one that is non-identical, even profoundly different in meaning and extension, so long as functional continuity is maintained. The needed functional continuity is, moreover, very flexible—functions can be rejected, traded off, split, combined, reshuffled.

We might call this 'radical' functionalism, in contrast to the 'moderate' functionalism suggested by the above accounts. While the moderate functionalist aims to answer Strawson's worries by arguing that we can preserve meaning or conceptual identity via preservation of function, the radical functionalist allows much more drastic revisions and outright replacements, holding only that our tinkering should not result in needed functions no longer being filled by some element of the post-engineering conceptual scheme.

Are there any radical functionalist proposals currently on offer in the extant conceptual engineering literature? A few functional accounts are at least consistent with radical functionalism, in that they do not explicitly commit themselves to the moderate functionalist goals of preservation of meaning or concept identity. Burgess and Plunkett (2013b) write of 'conceptual ethics' in a way that leaves open radical revision; for instance, they explicitly characterize eliminative materialism as a position concerning conceptual ethics. They also speak of our aims and goals as guiding conceptual revision, which makes them quite functionalist in flavor (though they also speak of 'goods' such as truth which they hesitate to reduce to goals). Thomasson (forthcoming) advocates a conception of conceptual function or purpose which is not couched in semantic terms; she suggests appealing to notions of function available in other literatures, such as biological notions of proper function. Haslanger (2000) provides what might be the closest to an explicit endorsement of a radical functionalist view. She encourages the engineer to ask what purposes are served by the target concepts; and she writes that new concepts to better serve these purposes may be introduced via pure stipulation. In later work, however, Haslanger appears to back away somewhat from the 'radical' aspect of her functionalism, writing that her approach does not change language so much as reveal "that our linguistic practices have changed in ways we may not have noticed" (Haslanger 2006, p. 106).

With the basic position now on the table, it's well past time to face the elephant in the room. I have defined the radical functionalist in a rather 'negative' manner, via what she does not do: she does not explain conceptual function in semantic terms, nor in terms of 'essential' features of concepts, and she does not prohibit dramatic, subject-changing revision. But I have not provided a positive account of the nature of the functions that constrain successful revision. This is by design; I haven't myself got much of a view on offer, nor do I think there's a clearly workable candidate in the existing literature. This apparent lacuna might seem problematic. Cappelen (2018),

Footnote 15 continued
human walking times. Successor concepts like 'kilometer' abandon that function, and aren't particularly worse off for doing so. 'Kilometer' certainly isn't the same concept as 'league'; it's a replacement rather than a revision.

for instance, rejects functional approaches to engineering precisely because he holds that there is no satisfactory explanation of what it might be for a concept to have a function; he rejects both Haslanger's and Thomasson's attempts to explain said function. But I'm rather unfazed. I doubt that the functionalist really needs the sort of account Cappelen demands; I'll argue that neutrality on the nature of function is here not only permissible, but appropriate.

Despite use of the 'function' label, talk of purposes, goals, or aims might be a bit more well-suited to a radical functionalist take on engineering. Cappelen dislikes functionalist accounts because he doesn't think concepts *have* functions. But humans certainly have purposes and goals and aims; and it's not particularly problematic to claim that we often use concepts to help us achieve those purposes, goals, and aims. *Contra* Thomasson, I don't think we ought to limit ourselves to speaking of something like the 'proper' function of a concept. Tools, concepts included, can be used for purposes other than their 'proper' function; I can use a hammer as a paperweight, or a knife as a lever. Weighting papers may not be the proper function of a hammer, but insofar as I use my hammer for that purpose, it makes perfect sense to speak of weighting papers as one of the hammer's (current) functions.

So suppose we decline to identify some central, or proper, or essential function of a concept—suppose instead we think of a concept's functions as being merely extrinsic, relational properties like 'is used by x for y'. Then the potential 'functions' of a concept will be deeply wide-ranging and various. And it becomes rather unlikely that a single account will unify all the myriad possible uses an engineer might deem relevant when considering the success of a revision. Indeed, even features that aren't comfortably termed 'functions' may affect our assessment of engineering success—ease of use or learnability, for instance.[16]

I think that, intuitively, all this messiness and inclusiveness is the right result. Is it really feasible, for instance, to demand a unified account of all the various 'functions' a product of physical re-engineering might need to preserve? Philosophical accounts of proper function, or systemic function, or any other sort of function might be applicable to human artifacts, but such accounts are too narrow to capture all that might be relevant to successful re-engineering of said artifacts. Insofar as a tool is being regularly used by a group of people for a certain purpose, regardless of whether that purpose is the tool's 'function' in the philosopher's sense, the continued fulfilment of that purpose might need to be considered when a proposed replacement is on offer. That's not to say that the function *must* be retained; we might find some other tool to serve the purpose, or we might decide that the purpose is no longer needed. We might even retain the older tool solely for the uses the replacement fails to adequately perform. But if a replacement tool can no longer fill one of the uses to which its predecessor

[16] Though it is odd to label such features 'functions', aim or goal language seems to do fine—e.g., we might design concepts to help us fulfil our goal of making a certain sort of classification practice less difficult. An example here might be the 'four food groups' that many Americans of a certain age learned in primary school; that rough-and-ready method of classifying foods vastly oversimplifies the diversity of their nutritional properties, but has the benefit of being easy to learn and deploy.

is put, this is *prima facie* a drawback. Appeal to such failures can fuel a legitimate criticism of the replacement.[17]

Thus, a radical functionalist concerned with continuity will need to ask: does our proposed replacement concept fill *all* the purposes to which the predecessor is put? If not, is this loss in function remedied or offset in any way (perhaps by addition of a further concept, or retention of the former for some contexts of use)? If not, is the loss in function justified (perhaps by argument that the function is no longer necessary, or that a trade-off for other functional utility is warranted)? This, admittedly, leaves us with a much larger task than merely retaining a 'core' function, or retaining sufficient similarity of meaning. But no one said conceptual engineering had to be easy.

Does this approach leave us beholden to analysis? Not in any strong way. If we don't attend carefully to the uses of our current concepts, we do certainly risk neglecting important functions in the design of our successor concepts. But several points are worth noting. First, some form of Craig-style conceptual ethnography seems more suited to the engineer's needs than standard conceptual analysis—perhaps informed by such empirical disciplines as anthropology and psychology (indeed, perhaps by experimental philosophy). Second, and relatedly, counterexamples produced by analysis will not support a legitimate criticism of a successor concept unless the critic can make a case that the counterexample reflects a neglected function. It's not enough to know *that* case x fails to fall under the pre-theoretic concept F; we must be told *why* the pre-theoretic concept excluded said cases, in the strong sense of being told what purpose that classificatory practice would be in service of.[18]

Finally, even if one discovers (via analysis or otherwise) a function which has been neglected in a proposed successor concept, this doesn't doom the successor. The engineer can in most cases simply propose a further new concept to take over the neglected function, or allow that the old concept should be retained in contexts where said function is needed. Or she may simply argue that the function is not worth retaining. The radical functionalist engineer, then, plausibly has much more flexibility and freedom than an engineer constrained by a semantic similarity desideratum. Yet she is constrained—despite Jackson's worries, philosophical problem-solving is not reduced to trivial stipulation.

[17] As an example: a microwave oven can be criticized as a wholesale replacement for a conventional oven on the grounds that, while a microwave cooks food more efficiently, it does not easily keep foods warm for long periods of time (as one might in the oven when mealtime is delayed).

[18] One response here, of course, could be along the lines of arguments given in Cappelen (2018): Cappelen claims that he sees no obvious function for e.g., "salmon" other than to talk about salmon. So we might worry that if our successor concept doesn't respect apparent counterexamples, then we are no longer talking about salmon and thus our successor is not fulfilling the concept's function. But even if we accept this move, functions can be critiqued (as noted below). So it is open to the functionalist to say—"I don't see why that function is worth retaining—why, that is, we should care about talking about salmon. I think it would be better to talk about shmalmon, for the following reason". We can assess the value of functions/purposes/aims, and indeed we can do so by appeal to further functions/purposes/aims. Just as in physical engineering: such-and-so structural element might function to increase the maneuverability of a vehicle, but it might also decrease speed, and whether that function is worth preserving might further depend on the use to which the vehicle will be put.

5 Truth and efficacy

So suppose I've convinced you that the proper continuity desideratum for a conceptual engineer is to be cast in functional terms, in roughly the way just outlined. There is an interesting, and likely to some unwelcome, corollary that I think must follow from such a view: if we take philosophy's primary goal to be functionalist conceptual engineering, then the primary aim of philosophy is neither truth nor knowledge. In fact, I suspect that truth (and consequently knowledge) is of only secondary value for *any* conceptual engineer, functionalist or otherwise. We look to truth when we evaluate sentences, propositions, beliefs; if a belief is true, it 'gets things right'. But the primary task of a conceptual engineer isn't evaluating sentences or beliefs. It's evaluating the concepts they are couched in. And a concept is not 'true' or 'false'. When we evaluate concepts, we are not looking for truth or falsity; we are looking for other sorts of virtues or flaws, such as exactness, or match with the joints of nature.

This suggests that conceptual engineers (functionalist or otherwise) need some analogue to 'true', some evaluative term that expresses the notion that a concept can 'get things right'. That analogue will be the primary measurement of success in conceptual engineering. Unfortunately, standard epistemological terminology doesn't give us a word that expresses 'getting things right' at the level of concepts. Carnapian explication might suggest that 'fruitfulness' would do the trick—but as Kitcher (2008) notes, fruitfulness in the sense of enabling universal generalizations is plausibly too narrow for even many of the sciences, not to mention the sorts of everyday uses that someone like Strawson has in mind.

One might think that a successful analogue in the same spirit could be something like 'accuracy' in matching the structure of reality—that is, we might claim that what determines the success of a concept is its fit with the 'joints of nature'. But I would argue that this is still too limiting. This is not to dismiss the idea that nature has joints—it's merely to claim that some concepts have different success conditions than accurately carving at said joints. For some purposes we might want reality thinly sliced, ground to mince, or cross-cut through the bone. In some cases, we don't even carve at all: consider useful but uninstantiated idealizations and abstractions in the sciences, such as the notion of a point particle, a frictionless plane, or an ideal gas.

The functional picture of conceptual engineering, however, presents a clear candidate—one which provides a broader notion of conceptual success. On the functional picture, a concept is successful if it effectively fulfils its functions—in short, if it is *effective*. We could say, then, that the goal of conceptual engineering is to design effective concepts—in the same way that a 'physical' engineer might aim to design an effective airplane wing, an effective juicer, an effective data storage system. In each case, success is measured in terms of how well the design fulfils the intended function. Success is measured in terms of *efficacy*.

Nothing rules out the notion that a concept might have, as its primary function, carving the world at its joints. Indeed, I think a great many concepts will have this function. So adopting a functional take on conceptual engineering in no way implies any form of metaphysical or scientific antirealism. It simply allows that in some cases, it is useful to have concepts which do something other than joint-carving. As Strawson points out, the 'fundamental' language with which we describe reality in the sciences

might be unsuited to many of the ordinary tasks to which language is put. There is a time and place for a category which, despite the joints of botany, groups tomatoes with vegetables rather than fruit. The place is the kitchen, and the time is when one is making a fruit salad. If conceptual success is put in terms of efficacy, we can measure success for those less 'fundamental' sorts of concepts too.

So we need not be anti-realists to embrace efficacy as the aim of conceptual engineering—and we need not be subjectivists in any sense, either, even for non-joint-carving concepts.[19] Making success a matter of fulfilling a desired function does not imply that there are no objective, mind-independent facts about which concepts are better and which are worse. There are, after all, objective, mind-independent facts about which computers are faster, which self-driving cars have lower accident rates, and so forth. There might even be (though I'm skeptical about this myself) objective facts about which functions ought to be filled. Finally, it should be clear that proposing efficacy as the aim of engineering is very much *not* to propose anything like pragmatism about truth or meaning. One can embrace a causal-historical view of reference, a correspondence view of truth, and a functional take on conceptual engineering without contradiction.

Though it is not inherently anti-realist or subjectivist, the notion that engineering aims at efficacy does makes success a bit more 'slippery' than would a rigid, all-or-nothing label like truth. Efficacy, unlike truth, is a matter of degree. With regard to airspeed, propeller planes are more effective than the Wright brothers' earliest, propless gliders—but jetplanes are more effective still. And as this same analogy makes clear, efficacy is also relative to a purpose or goal—a jetplane is more effective with regard to airspeed, but a helicopter is more effective with regard to maneuverability. Moreover, given the multiplicity of functions a design might pursue, increases in efficacy often involve trade-offs: durability increased at the cost of speed, or computational power increased at the cost of energy efficiency.

All this is easily applied to concepts, as well. The efficacy of a concept will be a matter of degree, and therefore improvement may be more-or-less. A successful conceptual analysis (if there ever were to be one) in a sense 'closes the door' on further inquiry—much as a proof of a theorem would. A successful instance of conceptual engineering, however, does not; a later philosopher may propose additional improvements. Since efficacy is relative to a goal, a concept may become 'obsolete' as our interests and goals change over time; and, of course, a proposed concept may be highly effective for one purpose but not another. And since our pretheoretic concepts are likely to serve multiple functions, engineered replacements may need to deal in trade-offs.

All of this fits very comfortably with the radical functionalist picture of conceptual engineering. We've noted that nothing prevents the radical functionalist from retaining an earlier concept for some purposes, or from proposing multiple successor concepts to fulfil functions that were previously filled (perhaps crudely) by a single concept—just as we might trade in our pocketknives in favor of a pair of scissors and a saw. Finally, we might note that it is at least an open possibility that any given function or set of

[19] Though a reviewer rightly points out that certain conceptual functions might as a matter of fact make the *particular* concepts that possess them inherently subjective (the reviewer notes, e.g., that one function of DELICIOUS might be to articulate subjective differences in taste).

functions could be fulfilled equally effectively by multiple, competing concepts—that there could be, in other words, efficacy 'ties'. Again, this does not imply that anything goes. Ties are compatible with objective facts about levels of efficacy.

So I would argue that a philosopher, if she is a conceptual engineer, should be happy to discover truths—but only if those truths couched in effective concepts. And it is the task of generating effective concepts, rather than the task of uncovering truths, which is primary. Contrast this with the goal of conceptual analysis, as traditionally understood. There, the concepts are fixed, and the goal is to generate an analysis—a *true* biconditional. But since a radical functionalist engineer is permitted to generate new concepts by pure stipulation, if she so desires, true biconditional analyses will be trivially easy to generate. Here, let's try—I hereby stipulate that a 'brollop' is a desk chair with five legs. Now, here is a truth: x is a brollop iff x is a desk chair with five legs. As conceptual engineers, we can have a grand old time generating truths in this fashion, but it's clear that the real challenge lies elsewhere. The challenge is to generate concepts that are effective, not analyses that are true.

This paper's primary topic has been conceptual engineering, but the contrasts just indicated between truth and conceptual success also raise a prima facie challenge for contemporary analytic epistemology. Epistemic success, for the standard epistemologist, is characterized by knowledge—that is, by some Gettier-proofed modification of the traditional 'justified true belief' formula. But one can have justified, true beliefs couched in all sorts of ineffective, even downright gruesome concepts. I *know* that the tree outside my window is grue. But this is much less valuable, epistemically speaking, than my knowledge that it is green. A move from a defective concept to an improved one is, plausibly, an epistemic improvement. When I make my concepts more precise, or make conceptual distinctions I did not previously possess, or generate a classificatory system that better matches with the joints of nature, I gain something—some epistemic good that goes beyond an increase in justification, or in true beliefs. Some epistemic good that contemporary epistemology largely neglects.

Of course, in many cases, the improvement of one's concepts will be the *result* of a gain in true beliefs—of coming to realize that whales are warm-blooded and lack gills, for instance. But one could have those beliefs and nonetheless retain a concept which groups whales and fish together. Of course, post-Putnam, the argument could be made that 'fish' is not (and never was) that concept. But one could certainly imagine a society that used a descriptively-referring concept fish* whose extension was as Carnap imagined. One could further imagine a society that did so despite knowing e.g., that whales are warm-blooded, and so forth. When a member of that society believes 'whales are fish*', he believes something true—and there's no reason to suppose he could not do so justifiedly, and thereby possess knowledge. But I'd argue that, were he and his society to shift to true beliefs couched in Piscis rather than fish*, this would be an epistemic gain.

As Ted Sider has recently put the point, "For a representation to be fully successful, truth is not enough; the representation must also use the right concepts" (Sider 2011, p. vii). If we aim for complete epistemic success, we should not merely aim to (justifiedly) believe that which is *true*; we must aim for true beliefs that are *couched in proper concepts*. I've plumped for efficacy, though admittedly, one concern about efficacy as the measure of conceptual success is that gains in efficacy are not always

clearly epistemic. Even if this is right, however, I'm not too bothered—I'm open to the idea that a narrower notion of conceptual success might be more appropriate for epistemology. But just as we might have non-epistemic reasons to make belief revisions (e.g., pragmatic or moral reasons), we might similarly have non-epistemic reasons to make conceptual revisions. I'd argue that the goals of conceptual engineering (and thus, to my eyes, of philosophy) encompass all such reasons—thus, while a narrower notion of conceptual success may serve best for epistemology, efficacy serves best for metaphilosophy.

Ultimately, even those wholly unconvinced by the functionalist approach and its attendant focus on efficacy should consider how 'conceptual success' fits into our overall epistemological picture. If one's epistemic state can be improved by improving one's concepts, even in cases where there is no accompanying improvement in the traditional tripartite features of knowledge, then perhaps we ought to subject 'knowledge' to a bit of conceptual engineering.

Acknowledgements Thanks to Max Deutsch, Michael Johnson and Dan Marshall for helpful input on earlier drafts of this paper. Thanks also to the very helpful comments from the reviewers for Synthese.

References

Alexander, J., & Weinberg, J. M. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass, 2*(1), 56–80.

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

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

Buckwalter, W., & Stich, S. (2014). Gender and philosophical intuition. In J. Knobe & S. Nichols (Eds.), *Experimental philosophy* (Vol. 2, pp. 307–346). Oxford: Oxford University Press.

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

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

Cappelen, H. (2012). *Philosophy without intuitions*. Oxford: Oxford University Press.

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

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

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

Carnap, R. (1963). Replies and systematic expositions. In P. A. Schilpp (Ed.), *The philosophy of Rudolf Carnap* (pp. 859–1013). La Salle: Open Court.

Craig, E. (1990). *Knowledge and the state of nature: An essay in conceptual synthesis*. Oxford: Clarendon Press.

Deutsch, M. (2009). Experimental philosophy and the theory of reference. *Mind and Language, 24*(4), 445–466.

Deutsch, M. (2010). Intuitions, counterexamples, and experimental philosophy. *Review of Philosophy and Psychology, 1*(3), 447–460.

Deutsch, M. (2015). *The myth of the intuitive*. Cambridge, MA: MIT Press.

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

Feltz, A., & Cokely, E. (2008). The fragmented folk: More evidence of stable individual differences in moral judgments and folk intuitions. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 1771–1776). Austin, TX: Cognitive Science Society.

Fisher, J. C. (2015). Pragmatic experimental philosophy. *Philosophical Psychology, 28*(3), 412–433.

Goldman, A. I. (1993). Epistemic folkways and scientific epistemology. *Philosophical Issues, 3,* 271–285.
Haslanger, S. (2000). Gender and race: (What) are they? (What) Do we want them to be? *Noûs, 34*(1), 31–55.
Haslanger, S. (2006). Philosophical analysis and social kinds. *Proceedings of the Aristotelian Society, Supplementary Volumes, 80,* 89–118.
Jackson, F. (1998). *From metaphysics to ethics: A defence of conceptual analysis*. Oxford: Oxford University Press.
Justus, J. (2012). Carnap on concept determination: Methodology for philosophy of science. *European Journal for Philosophy of Science, 2*(2), 161–179.
Kemeny, J. G., & Oppenheim, P. (1952). Degree of factual support. *Philosophy of Science, 19,* 307–324.
Kitcher, P. (2008). Carnap and the caterpillar. *Philosophical Topics, 36*(1), 111–127.
Laurence, S., & Margolis, E. (1999). Concepts and cognitive science. In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings*. Cambridge, MA: MIT Press.
Machery, E. (2017). *Philosophy within its proper bounds*. New York: Oxford University Press.
Machery, E., Mallon, R., Nichols, S., & Stich, S. (2004). Semantics, cross-cultural style. *Cognition, 92,* B1–B12.
Maher, P. (2007). Explication defended. *Studia Logica, 86*(2), 331–341.
Novaes, C. D., & Reck, E. (2017). Carnapian explication, formalisms as cognitive tools, and the paradox of adequate formalization. *Synthese, 194*(1), 195–215.
Olsson, E. J. (2015). Gettier and the method of explication: A 60 year old solution to a 50 year old problem. *Philosophical Studies, 172*(1), 57–72.
Pinder, M. (2017a). Does experimental philosophy have a role to play in Carnapian explication? *Ratio, 30*(4), 443–461.
Pinder, M. (2017b). On Strawson's critique of explication as a method in philosophy. *Synthese*. https://doi.org/10.1007/s11229-017-1614-6.
Plunkett, D. (2015). Which concepts should we use? Metalinguistic negotiations and the methodology of philosophy. *Inquiry, 58*(7–8), 828–874.
Plunkett, D., & Sundell, T. (2013). Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint, 13*(23), 1–37.
Prinzing, M. (2018). The revisionist's rubric: Conceptual engineering and the discontinuity objection. *Inquiry, 61*(8), 854–880.
Quine, W. V. O. (1960). *Word and object*. Cambridge: MIT press.
Scharp, K. (2013). *Replacing truth*. Oxford: Oxford University Press.
Schupbach, J. (2015). Experimental explication. *Philosophy and Phenomenological Research, 94*(3), 672–710.
Shepherd, J., & Justus, J. (2015). X-Phi and Carnapian explication. *Erkenntnis, 80*(2), 381–402.
Sider, T. (2011). *Writing the book of the world*. Oxford: Oxford University Press.
Strawson, P. (1963). Carnap's views on conceptual systems versus natural languages in analytic philosophy. In P. Schlipp (Ed.), *The philosophy of Rudolf Carnap* (pp. 503–518). La Salle: Open Court.
Swain, S., Alexander, J., & Weinberg, J. (2008). The instability of philosophical intuitions: Running hot and cold on Truetemp. *Philosophy and Phenomenological Research, 76*(1), 138–155.
Thomasson, A. L. (2017). What can we do, when we do metaphysics? In G. D'Oro & S. Overgaard (Eds.), *Cambridge companion to philosophical methodology* (pp. 101–121). Cambridge: Cambridge University Press.
Thomasson, A. L. (Forthcoming). A pragmatic method for conceptual ethics. In Cappelen, H., Plunkett, D., & Burgess, A. (Eds.), *Conceptual ethics and conceptual engineering*. Oxford University Press.
Tobia, K., Chapman, G., & Stich, S. (2013). Cleanliness is next to morality, even for philosophers. *Journal of Consciousness Studies, 20*(11–12), 195–204.
Weinberg, J., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29,* 429–460.