DOI: 10.1111/nous.12538

ARTICLE

Noûs

Heavy-duty conceptual engineering

Steffen Koch[1] | **Jakob Ohlhorst**[2]

[1]Bielefeld University, Bielefeld, Germany

[2]Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Correspondence
Jakob Ohlhorst, Vrije Universiteit Amsterdam,, Amsterdam, Netherlands.
Email: j.ohlhorst2@vu.nl

Both authors have contributed equally to this work; they are joint first authors

Funding information
Alexander von Humboldt-Stiftung; DEU 1225691 FLF-P; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung; P500PH_211167

Abstract

Conceptual engineering is the process of assessing and improving our conceptual repertoire. Some authors have claimed that introducing or revising concepts through conceptual engineering can go as far as expanding the realm of thinkable thoughts and thus enable us to form beliefs, hypotheses, wishes, or desires that we are currently unable to form. If true, this would allow conceptual engineers to contribute to solving *stubborn problems* – problems that cannot be solved with our current ways of thinking. We call this kind of conceptual engineering *heavy-duty conceptual engineering*. As exciting as the idea of heavy-duty conceptual engineering sounds, it has never been developed or defended. In this paper, we pursue a two-fold goal. First, to offer a theory of heavy-duty conceptual engineering that distinguishes it from other kinds of conceptual engineering; second, to show that heavy-duty conceptual engineering is possible, both in theory and in practice, and to explain how it can be applied in the service of solving stubborn problems. The central idea is that heavy-duty conceptual engineering can enhance the semantic expressive power of a conceptual system by the use of bootstrapping processes.

1 | INTRODUCTION

Conceptual engineering is the process of assessing and improving our conceptual repertoire. Many philosophers have recently argued that conceptual engineering is and should be practiced in philosophy, the sciences and beyond.[1] But what can we hope to achieve by doing conceptual engineering? In other words, why does conceptual engineering matter? This paper explores a radical answer to this question: that one of the benefits of conceptual engineering is that it can, quite literally, expand the realm of thinkable thoughts. Conceptual engineering enables us to form thoughts, including beliefs, desires, hypotheses, or wishes, that were impossible for us to form before.

The idea of expanding the realm of thought through conceptual engineering has never been seriously explored in the literature, but it has fueled its rhetoric nonetheless. An often-cited passage from Burgess and Plunkett reads:

> "our conceptual repertoire determines not only *what beliefs we can have* but also *what hypotheses we can entertain*, *what desires we can form*, *what plans we can make* on the basis of such mental states, and accordingly constrains what we can hope to accomplish in the world." (Burgess and Plunkett, 2013, p. 1096f., our emphasis)

A similar thought is expressed when Sterken writes that introducing new concepts through conceptual engineering may "enable an interpreter to think and communicate things she could not have thought or said without having that meaning" (Sterken, 2020, p. 430); or when Kevin Scharp argues that conceptual engineering can be a means to "take some control over what we can *think, say, do, and be*" (Scharp, 2020, p. 398, our emphasis). All of these passages contain versions of the core idea of this paper: conceptual engineering can provide us with access to previously unthinkable thoughts.

In the engineering sciences, "heavy-duty engineering" refers to manufacturing raw materials, whereas "light engineering" turns pre-fabricated goods into saleable products. Analogously, we will use the label of *heavy-duty conceptual engineering* to refer to the creation of new "raw materials of thought" which can be used to form new kinds of thoughts.[2]

The idea of heavy-duty conceptual engineering is exciting, as it suggests a way in which philosophical and other forms of theorizing can generate entirely new ways of thinking. Conceptual engineering is often characterized as a problem-solving method (cf. Isaac et al., 2022). Arguably, some of the theoretical and practical problems that we face cannot be solved by our current ways of thinking. We call these *stubborn problems*.[3] By giving us access to previously unthinkable thoughts, heavy-duty conceptual engineering may offer a method for solving such stubborn problems.

But the idea is also controversial. One problem is that it is unclear what "making previously unthinkable thoughts thinkable" could possibly mean, let alone how it could be done by conceptual engineers. Some philosophers even doubt that conceptual engineers are in the business of introducing new concepts to begin with. According to Matti Eklund, for example, conceptual engineering "concerns which concepts we more easily or readily *employ*, and not which concepts we *have*" (Eklund, 2021, p. 18, our emphasis). If this is right, then conceptual engineering serves to

[1] See Isaac et al. (2022) and Koch et al. (2023) for recent overviews.

[2] Scharp (2013) also uses the phrase "heavy-duty conceptual engineering", but in a different sense.

[3] Consider, for example, how thinking in terms of Euclidean space did not allow us to solve the problems of spacetime relativity.

change what we do with the concepts we already have instead of providing us with genuinely novel ones. Eklund therefore holds that the rhetoric about the mind-expanding powers of conceptual engineering “should [...] be tempered” (ibid.).

While previous discussions of conceptual engineering have touched upon many difficulties with designing and especially implementing new or revised concepts, as of yet, there is no account specifically concerned with heavy-duty conceptual engineering and the problems that it gives rise to. The present paper takes a first step towards closing this gap. It aims to answer the following questions: What does heavy-duty conceptual engineering amount to? What does it mean for a thought to be “unthinkable”? How can previously unthinkable thoughts become thinkable? Is heavy-duty conceptual engineering possible in practice? And if so, how can we use it to solve stubborn problems? In answering these questions, this paper pursues a twofold goal: first, to offer a general framework for how we can make sense of heavy-duty conceptual engineering and how it can be distinguished from other kinds of conceptual engineering; second, to show that heavy-duty conceptual engineering is possible in practice and how it can be done. It will emerge that Burgess and Plunkett’s optimism about the potency of conceptual engineering is partly justified: heavy-duty conceptual engineering is indeed possible, both in theory and in practice. But Eklund’s plea for sobriety is also partly justified, as many of the frequently cited paradigm cases of conceptual engineering in philosophy do not qualify as heavy-duty.

The plan is as follows. In section 2, we begin by underpinning the general idea of heavy-duty conceptual engineering with a theoretical framework. We will suggest that heavy-duty conceptual engineering is a kind of conceptual engineering that alters a community’s semantic expressive power through the introduction of collectively foreign concepts. In section 3, we supplement this framework with a mechanism of introducing and learning previously foreign concepts – *boot-strapping*. In section 4, we explain how heavy-duty conceptual engineering can be used to tackle stubborn problems. In section 5, we consider two important objections against the framework.

Before we begin, three clarifications are in order. First, there is currently no consensus about what conceptual engineers are in the business of engineering, e.g. linguistic entities such as semantic meanings (Cappelen, 2018) or speaker meanings (Pinder, 2021), mental representations such as concepts (Isaac, 2023), or yet other things like classification procedures (Nado, 2023). The focus of this paper is set entirely on mental representations; more specifically, on thoughts and concepts, understood as abstract types with concrete mental tokenings. Still, as we will see, language plays an important role in heavy-duty conceptual engineering.

Second, the conceptual engineering literature typically distinguishes *conceptual innovation* (Simion & Kelp, 2020) or *de novo engineering* (Chalmers, 2020) from *conceptual re-engineering* (Brun, 2016). As Simion and Kelp (2020) argue, there is no reason to restrict the engineering endeavor to fixing the defects of existing concepts, instead of also seeking to innovate new ones. Following them, we will understand conceptual engineering in a broad sense, so as to include both conceptual innovation and re-engineering. Since the central question of this paper is whether a conceptual system can be transformed such that it yields access to previously unthinkable thoughts, it won’t matter whether this is done by conceptual re-engineering or innovation.

Third, conceptual engineering evaluates concepts along, e.g., moral (Haslanger, 2012) or epistemic dimensions (Egré & O’Madagain, 2019). In this paper, however, we will sidestep normative issues about what makes concepts fruitful (Carnap, 1950; cf. Pinder, 2020) or deficient (Cappelen, 2018), or what grants them authority (Queloz, 2022). Instead, our goal is to make sense of the general idea of conceptual transformations as described above. For all we say in this paper, changes can be for better or for worse. We leave it to future research to supplement our proposal with an account of when heavy-duty conceptual engineering is also *ameliorative*.

2 | WHAT IS HEAVY-DUTY CONCEPTUAL ENGINEERING?

2.1 | Building blocks, expressive power, and foreign concepts

A natural starting point for theorizing about heavy-duty conceptual engineering is what we might call the *building blocks view* of concepts – the view that concepts are the building blocks (or constituents, or units) of thoughts and propositions. The building blocks view is popular among leading contemporary philosophers of mind (Rey, 1983; Fodor, 1998; Carey, 2009; Margolis & Laurence, 2019). It is naturally coupled with a view about *possession* and *thinkability*: for a subject to be able to think a thought (or to entertain a proposition), she needs to possess the concepts from which it is built (*cf*. Burgess & Plunkett, 2013, p. 1095; Simion & Kelp, 2020, p. 986), and for her to possess a concept, she needs to be able to deploy it in thought.

We will soon see that this idea is in need of refinement. Nonetheless, we can already see how it vindicates the core idea of heavy-duty conceptual engineering: if the realm of thinkable thoughts is constrained by the concepts we possess, then we can expand the realm of thinkable thoughts by acquiring new concepts, or by revising those we possess already. As conceptual engineering is a means of generating new or revised concepts, conceptual engineering can, at least in principle, expand the realm of the thinkable.

However, it is clear that not every conceptual innovation yields access to previously unthinkable thoughts. Consider, for example, the following stipulations: LESK = large desk; BLALL = blue wall; RINDOW = window that opens to the right side. You have probably never heard of lesks, blalls or rindows before. But did reading these stipulations enable you to think previously unthinkable thoughts, thoughts that feature lesks, blalls or rindows respectively? We take it as uncontroversial that this is not so. This is because concepts like LESK, BLALL and RINDOW do not add anything genuinely new to your overall conceptual repertoire. But why not? What does it take for a concept to add something genuinely new to a given conceptual system?

This question is broadly debated among philosophers of mind (Fodor, 1998; Carey, 2009; Rey, 2014; Beck, 2017). Various attempts have been made to clarify the difference between the kind of examples above (LESK, BLALL, RINDOW) and genuinely new conceptual additions. According to an influential proposal by Carey (2009), the crucial question is whether the concept in question is *continuous* or *discontinuous* with the previous set of concepts, where a concept is discontinuous with a pre-existing conceptual system just in case the former cannot be defined in terms of the latter, or be translated into it. Similarly, Rey (2014) introduces the notion of *expressive power*, distinguishing between concepts that can be formed by logical combinations of already existing concepts, and those that cannot. The former concepts are *possessed* but not *manifested*, whereas the latter are not even possessed. Bringing all of these pieces together, Beck (2017) distinguishes between three types of concepts and two forms of expressive power. In what follows, we will briefly sketch the central idea behind these distinctions, before using them to define different varieties of conceptual engineering.

A good starting point is what Beck calls an *available* concept. According to Beck, "a concept is available to a person just in case she could deploy it in reasoning, thinking, categorizing, remembering, and other cognitive processes without much effort, simply by endogenously shifting her attention" (Beck, 2017, p. 113).[4] Beck emphasizes that many more concepts are available to you than those you are deploying at a given time. While you might not have thought of a mouse until

[4] Beck's available concepts match Rey's manifested concepts – concepts that are "actually or readily activated" (Rey, 2014, p. 112).

a second ago, clearly the concept MOUSE was available to you even then. Thinkers may use their available concepts to form further complex concepts and propositions. The *psychological expressive power* of a subject comprises all the complex concepts and propositions that this subject is psychologically able to build from their available concepts (Rey, 2014, p. 127; Beck, 2017, p. 114).[5]

Available concepts can be distinguished from latent (Beck, 2017) or possessed yet unmanifested (Rey, 2014) concepts. One way for a concept to be latent is by being composed from available concepts in an unobvious fashion, i.e., in such a way that the resulting concept cannot be easily deployed in reasoning, thinking, categorizing, etc. due to limitations in our psychological capacities.[6] The distinction between available and latent concepts is not sharp, but there are clear instances on both sides. On the one hand, the stipulated concepts LESK, BLALL and RINDOW are complex concepts that are composed from available concepts, but they are so straightforward that they will not count as latent in the sense of unobvious composition. Sally Haslanger's revisionary gender concepts (Haslanger, 2012, p. 234), on the other hand, are likely to be latent to those who lack sufficient training with them (see below). Even though they are composed of concepts that are available to most people, they are so complicated and deviate so much from ordinary gender classifications that they are quite hard to deploy. The *semantic expressive power* of a subject comprises all the complex concepts and propositions that can be built by an ideally rational agent from combinations of her available and latent concepts combined (Rey, 2014, p. 127; Beck, 2017, p. 114).[7]

Lastly, both available and latent concepts can be distinguished from *foreign* concepts. Foreign concepts are neither available nor latent. A foreign concept is discontinuous from the subject's conceptual system, while available and latent concepts are continuous with the individual's conceptual system. It is controversial whether there really are foreign concepts, or whether all concepts are either available or latent, as nativists would argue (see section 5 for discussion). For illustration, and assuming that there are foreign concepts to begin with, consider what the concept of SPACETIME is for five-year old Leah. It is neither available for her nor can she compose it from the concepts available to her. She also does not have the concepts of ELECTROMAGNETIC FIELD, GRAVITATION, or the prerequisite mathematical concepts available, nor could these concepts be composed from her available concepts. By acquiring one of these concepts, Leah would increase both her psychological and her semantic expressive power.[8]

The notions of semantic expressive power and foreign concepts also allows us to clarify what makes problems stubborn: a problem is stubborn for us if there is no solution that lies within our current semantic expressive power. To the extent that a stubborn problem can be solved at all, it can only be solved by introducing formerly foreign concepts.

2.2 | Conceptual engineering: from light to heavy-duty

Concepts are not available or foreign per se, but relative to particular subjects and times. The same holds for the related notions of expressive power. However, conceptual engineering typically goes beyond individual concept learning and targets the conceptual repertoire of entire

[5] Psychological ability is limited by aspects like computing power, memory constraints, emotional blockades, etc.

[6] Rey (2014) and Beck (2017) argue that latent concepts can also be innate but unactivated primitive concepts. While we agree that there are innate concepts, namely *core cognition* (Carey, 2009), they are quite limited.

[7] Here and elsewhere, we follow the mainstream view in understanding ideally rational agents as agents who possess omniscience and infallibility in logical but not in empirical domains (cf. Christensen, 2004; Smithies, 2015).

[8] See section 5 for objections against the existence of foreign concepts.

epistemic communities. To capture this important aspect of conceptual engineering, we separate a single *individual's* expressive power from an entire *epistemic community's* expressive power. A distinction along these lines plays a crucial role in the literature on conceptual change and innovation as it marks the difference between mere *concept acquisition* and *conceptual innovation* (e.g. Boden, 1990, Nersessian, 2008). Drawing this distinction in precise terms is difficult because it connects with broader issues in metaphysics, social epistemology, and the philosophy of mind. But the basic idea is quite simple: In the course of their lives, individuals need to acquire many concepts that they were not born with, but for whose acquisition they can rely on their community to expand their individual expressive power. Meanwhile, some concepts are foreign to everyone – no member of the community possesses them, and hence the community does not possess them. In this case, individuals or groups within the community need to innovate the foreign concepts in order to expand the expressive power of the community. Bearing this in mind, we suggest distinguishing between the following three types of conceptual engineering:

> *Light conceptual engineering*: A type of conceptual engineering that alters neither a community's psychological nor its semantic expressive power by introducing new labels for concepts that are already available.
>
> *Moderate conceptual engineering*: A type of conceptual engineering that alters a community's psychological expressive power by constructing unobviously composed complex concepts out of already available concepts.
>
> *Heavy-duty conceptual engineering*: A type of conceptual engineering that alters a community's semantic expressive power by introducing collectively foreign concepts.

Since the distinction between available and latent concepts is not sharp, neither is the distinction between light and moderate conceptual engineering. There are nevertheless some reasonably clear examples for each of them.

In his recent book *Reality+*, David Chalmers recalls when he coined the term "hard problem of consciousness" in a talk in 1994:

> This name caught on more quickly than anything else I've ever said [...] [T]his was [not] because the idea was radical or original—in fact quite the opposite. The name caught on so fast because everyone knew what the hard problem was all along. (Chalmers, 2022, p. 278f.)

If we take his word for it, the coining of "the hard problem of consciousness" amounted to an instance of light conceptual engineering: people had thought about this problem for a long time and the phrase "the hard problem" was merely a catchy label for it. This episode further reveals that the word "light" in "light conceptual engineering" is value-free: episodes of light conceptual engineering can be momentous – introducing terms for previously unlabeled things often has significant cognitive and real-worldly effects (cf. Ritchie, forthcoming; Koch & Lupyan, 2024). The same holds for moderate conceptual engineering.

As an example of moderate conceptual engineering, we suggest Haslanger's revisionary definition of WOMAN, of which we will give you only the first and simplest of three clauses:

S is a woman iff

i) S is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a female's biological role in reproduction; [...] (Haslanger, 2012, p. 234)

Even this partial definition gives a lot to chew on, despite the fact that most competent English speakers will know the words and concepts that feature on the right side of the biconditional. For this reason, the concept defined by Haslanger serves as a paradigmatic example of a latent concept: a complex concept whose components are typically available but are arranged in such an unobvious way that it is difficult to use. Making latent concepts available augments one's psychological expressive power. A type of conceptual engineering that has this effect on an epistemic community classifies as moderate conceptual engineering.[9]

Whether heavy-duty conceptual engineering is possible – both in theory and in practice – is controversial and requires separate argument. Despite the common rhetoric that conceptual engineering enables us to think previously unthinkable thoughts, an explanation of how this could be done has not yet been presented in the literature. To be sure, there is an extensive debate about the possibility and feasibility of *moderate* conceptual engineering, with positions ranging from extreme pessimism to extreme optimism. Simion and Kelp (2020), for example, argue that the innovation and implementation of new concepts is readily feasible through changes in our habitat. However, these debates are primarily concerned with difficulties in the *uptake* of concepts. In contrast, the difficulties of heavy-duty conceptual engineering arise already at its *design stage*, for this requires the introduction of foreign concepts that cannot simply be stipulatively defined or produced by changes to our habitat. For the same reasons, the implementation of heavy-duty conceptual engineering is no less difficult, as we will discuss in section 4.

Before moving on to the task of showing that and how heavy-duty conceptual engineering is possible, we would like to suggest a refined notion of thinkability. In the beginning of this section, we described a connection between concept possession and the thinkability of a thought. This idea bears a straightforward connection to available and foreign concepts: if a given thought is composed entirely of concepts that are available to you, then it is thinkable for you; if it features foreign concepts, it is not thinkable for you. But what if it features latent concepts? The distinction between latent and foreign concepts requires disambiguating the modality of think*able*.

We argue that thinkability derives from our expressive power. Consequently, thoughts are either *unthinkable* because they lie outside our semantic expressive power, *semantically thinkable* because they lie within our semantic expressive power, or *psychologically thinkable* because they lie within our psychological expressive power. While moderate conceptual engineering can expand a community's realm of psychologically thinkable thoughts by making formerly latent concepts available, only heavy-duty conceptual engineering has the potency to alter a community's semantic thinkability by making formerly foreign concepts available.[10]

[9] The moderate/heavy-duty distinction stands orthogonally to the question how difficult it might be to use a newly engineered concept. For instance, Scharp's (2013) new concepts of truth are easily definable, yet it is difficult to implement them into logical reasoning or practical use.

[10] Note that, strictly speaking, the distinction between psychological and semantic thinkability also translates into a distinction between two types of stubborn problems: those whose solutions lie outside our psychological expressive power but inside our semantic expressive power, and those whose solutions lie outside of both. We take "stubborn problems" to refer to the latter only.

3 BOOTSTRAPPING

How do we introduce foreign concepts into a conceptual system if we are unable to define them? Here we suggest one empirically based model of how foreign concepts can be acquired, while leaving open the possibility that there may be others. Nersessian (2008) and Carey (2009) argue that children and scientists acquire foreign concepts through a process called *bootstrapping*. In this section, we present an account of bootstrapping inspired by their models.

We propose to distinguish between two kinds of bootstrapping that are often conflated in the literature on the topic: *educative bootstrapping*, an externally guided process that occurs in learning contexts, and *heuristic bootstrapping*, a creative process that occurs in research contexts. Educative bootstrapping occurs when individuals are taught to acquire *individually* foreign concepts that are already possessed by other members of their community. Heuristic bootstrapping occurs when scientists and researchers innovate *collectively* foreign conceptual frameworks that nobody in their community possesses yet. As we will see, both kinds of bootstrapping play an important role in heavy-duty conceptual engineering, although we will put greater emphasis on the second.

Before we explain bootstrapping in more detail, we want to briefly address an alternative and relatively easy way to acquire foreign concepts. As Carey (2009) and others have argued, humans possess a range of *domain specific learning mechanisms* (DSLMs) that generate foreign concepts upon the appropriate input.[11] Most notably, the DSLM of *psychological essentialism* gives access to previously foreign natural kind concepts: if we encounter *a* single instance of an unknown animal – e.g. a chevrotain – we are automatically disposed to generate a new concept of THIS KIND OF ANIMAL. While we see no principled objections to relying on DSLMs in the pursuit of heavy-duty conceptual engineering, we submit that this procedure is limited in three important respects: (i) It only works in cases where the subject has access to concrete and observable tokenings of a given kind, which makes it unfit for the introduction of theoretical concepts such as MASS, GRAVITATION, FREEDOM, or KNOWLEDGE. (ii) The use of DSLMs is an automated process that is not in the volitional control of the subject; the subject may not choose whether or not she acquires a concept in this way, nor what will be in the extension of the concept. And, relatedly, (iii) the use of DSLMs to acquire new concepts is not sensitive to normative reflection. If appropriately triggered, DSLMs will generate a given concept regardless of whether this would be a good concept to have. For these reasons, DSLMs and similar mechanisms are not the primary source for heavy-duty conceptual engineering, despite their capacity to deliver foreign concepts.

Let us now turn to bootstrapping. We will first present two case studies of educative as well as heuristic bootstrapping and then offer a general account of bootstrapping on their basis.

3.1 Educative bootstrapping

The process of educative bootstrapping and the related cognitive and conceptual issues are nicely illustrated by Carey's discussion of how children learn to count. Children initially lack any concept of natural number: neither individual natural numbers nor the abstract concept of NATURAL NUMBER are expressible with their conceptual repertoire. To them, number concepts are foreign. Nevertheless, toddlers are not a blank slate (Carey, 2009, p. 291f.), for they are

[11] Differently from Carey, Rey (2014) and Beck (2017) argue that these DSLMs are not mechanisms for the acquisition of foreign concepts but rather innate, i.e. potential or latent, concepts that lie in wait for a special kind of input in order to be activated.

equipped with innate cognitive resources that provide them with access to some very basic numerical and quantitative concepts.

We will simplify the picture here and assume that the only numerical conceptual resource that toddlers originally possess is *parallel individuation*.[12] This designates the capacity to keep track of up to three (sometimes four) objects at the same time. This capacity enables a toddler to know that these TWO OBJECTS are more objects than that ONE OBJECT. The child is also abstractly aware of this relation: even when the objects are hidden from her, she knows where there are more. Toddlers thus possess the rudimentary quantitative concepts of ONE OBJECT, TWO OBJECTS, and THREE OBJECTS as well as the relations of MORE and LESS. This is the original conceptual system *CS1* relative to which the concepts of natural numbers are foreign – they cannot be formulated with the child's original conceptual resources.

Children who learn how to count are not aware that their *CS1* cannot be used to think about numbers because possessing natural number concepts requires the concept of a recursive successor function that connects them. The learning process begins with learning to recite the count list: "one, two, three, …, ten". Initially, these words are entirely meaningless to the child. Even so, there is a clear order to the count list: "one" comes first, "two" comes next, and so on. So even though these number words are little more than meaningless placeholders, they are nevertheless structured (Carey, 2009, p. 308).

Next, reciting the count list is enriched with the practice of pointing to a different object at each word: *pointing to object a* "one", *pointing to object b* "two", *pointing to object c* "three". This establishes a *partial interpretation* of the count list: each word on the list relates to a different object, and the count list stops when there are no more objects. However, this is still insufficient to grasp how numbers work. It is little more than a game for the child, and the meaningless placeholders "one, two, three, …" only refer to a position within the counting game, not yet to natural numbers. Recall the toddler's pre-existing conceptual resources, *CS1*: she has the concepts of ONE OBJECT, TWO OBJECTS, and THREE OBJECTS. With time, she grasps that when there is only one object, the count list stops at "one" and that "one" means that there is only one object – she becomes a "one-knower". Later, she does the same with "two", and then with "three" – becoming a "two-" and then a "three-knower" (Carey, 2009, p. 320). She has mapped her notion of THREE OBJECTS onto the word "three". Consequently, the count list is now partially interpreted relative to *CS1*.

The connections between the concepts ONE OBJECT, TWO OBJECTS, THREE OBJECTS and reciting "one", "two", "three" constitute a set of constraints. In particular, the placeholder count list continues: "four, five, six…". By inductive reasoning as a modeling process the toddler figures out that each counting step adds one more object. Hence, "four" must be THREE OBJECTS AND ONE MORE and this rule also applies to every subsequent step, i.e. "five", "six", etc. Once toddlers can count past three, they can count as far as the numbers they know the words for. They have acquired the new conceptual system *CS2*. They do not have to work out each further step. Understanding the *n+1* rule through induction gives the child access to particular natural number concepts like ONE, FIVE, and EIGHT as well as the abstract concept of NATURAL NUMBER. The meaningless placeholder structure is now interpreted by applying it to quantities (see figure 1).

This is a simplified reconstruction of how children learn to count and acquire the natural number concepts, but it illustrates how educative bootstrapping works. The processes we just described are *educative*: they occur when children acquire concepts foreign to them but known to their community. But heavy-duty conceptual engineering aims at developing concepts foreign

[12] There is also a capacity to recognise proportions whose role we bracket here for simplicity's sake.

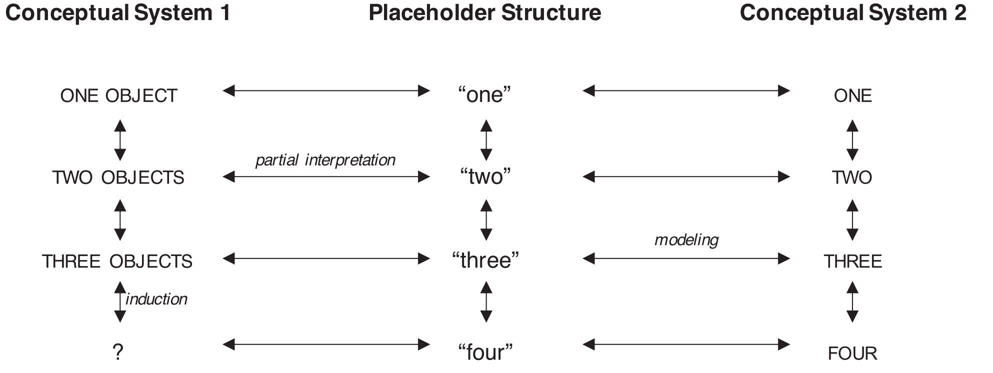

FIGURE 1 Bootstrapping natural numbers

to the entire community. How can this work without external guidance? This leads us to *heuristic* bootstrapping.

3.2 | Heuristic bootstrapping

Many philosophers of science argue that scientists innovate foreign concepts in scientific practice (Thagard, 1992; Kuhn, 1996; Gentner, 2002; Andersen et al., 2006; Nersessian, 2008; Carey, 2009). A prominent illustration is the case of Johannes Kepler who introduced the precursor concept to GRAVITATIONAL FORCE. The original *CS1* – the Ptolemaic conception of the solar system as a set of crystalline spheres that directly move the planets embedded in them – became inadequate due to a range of observations. *CS1* gave way to a new conceptual system, *CS2*, of planets being moved by FORCE AT A DISTANCE. The two systems were deeply discontinuous: in *CS1*, planets were moved by direct material contact with the crystalline spheres. *CS1* could not be recomposed into *CS2*, which required the concept of physical force being exerted at great distances on planets moving through empty space. To illustrate how foreign this concept of FORCE AT A DISTANCE was, consider what Newton, in a letter to Bentley, wrote about it:

> "It is inconceivable that inanimate Matter should, without the Mediation of something else, which is not material, operate upon, and affect other matter without mutual Contact… that one body may act upon another at a distance thro' a Vacuum, without the Mediation of any thing else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity that I believe no Man who has in philosophical Matters a competent Faculty of thinking can ever fall into it." (from Berkovitz, 2016)[13]

We will illustrate heuristic bootstrapping with a simplified version of Gentner's (2002) and Carey's (2009, pp. 422–428) reconstruction of Kepler's endeavor. *CS1*, the theory of rotating crystalline spheres was shattered through observations of comets passing through these spheres' locations. This generated a stubborn problem – what force could move and guide these bodies on their course? To answer this question, Kepler first replaced the crystalline spheres with a par-

[13] Similar sentiments can be ascribed to Lord Kelvin (see Quine & Ullian, 1978, p. 114).

tially interpreted placeholder structure of animate planets that were self-propelled and kept on course by their souls. This allowed him to freely substitute any element of this placeholder structure with other placeholders or features. The observation that the planets slow down the further they are from the sun imposed a constraint on the placeholder structure. For that reason, Kepler changed the placeholder structure to one where only the Sun possesses a soul that moves the planets – growing distance from the sun explains its force getting weaker (Gentner, 2002, p. 34). Kepler noted that light, a purely physical force, behaves analogously to this placeholder structure of the Sun's soul, exerting a force on the planets, both getting weaker with the square of the distance, and so on. He used this analogy and constraint to further transform his placeholder structure into a purely physical force exerted by the Sun on the planets. Kepler gave meaning to the placeholder structure of the Sun's soul moving the planets by interpreting it in purely physical terms analogous to light.

The episode of Kepler is just one of many examples of heuristic bootstrapping. Nersessian (2008) has reconstructed the complex bootstrapping processes involved in Maxwell's introduction of the foreign concept of the ELECTROMAGNETIC FIELD. Further plausible candidates include Mendel's concept of GENE and Einstein's concept of SPACETIME.

3.3 | Bootstrapping: A generalized picture

We can now develop a general account of bootstrapping. Often, what starts the process is that the subject is confronted with a *stubborn problem* – be it a phenomenon that cannot be adequately described with *CS1* (e.g., planetary movements), or a task that *CS1* does not allow them to perform (e.g., counting quantities). Bootstrapping helps cognizers to solve this problem by giving them access to concepts that are foreign to their original *CS1*, e.g., natural number concepts or the concept of FORCE AT A DISTANCE.

Bootstrapping is done in two crucial steps. In step one, the bootstrapper needs to introduce or learn an *uninterpreted* or *partially interpreted placeholder structure*. An *uninterpreted placeholder structure* consists of a set of completely meaningless placeholders that are stably connected by a set of formal relations; for example, an ordering, a lattice, or a set of graphs. It is these connections that turn the set of placeholders into a placeholder *structure*. Dutilh Novaes calls the reliance on these meaningless structures *de-semantification* (Dutilh Novaes 2012, p. 199). A *partially interpreted placeholder structure* adds interpretations to some of the placeholders, turning them into symbols that have some meaning. As we have seen in the Kepler case, placeholder structures derived from phenomena that are structurally analogous to the problem at hand are often promising. Sometimes one may also have to introduce a placeholder structure that directly maps onto structural features of the problematic phenomenon. The formal connections of the placeholders and their eventual partial interpretation impose *constraints* on the kinds of inferences and operations that the subject can perform with these symbols. The connections and interpretations need to be preserved, or, if they are changed, they need to be changed systematically to preserve as much of the broader structure as possible.[14]

[14] Carey (2014) discusses, for example, the role of constraints on induction. These constraints are related to, but go beyond, the debiasing effect of de-semantified formal structures which consists in bracketing prior beliefs from influencing a reasoning process (Dutilh Novaes, 2012, p. 222). They go beyond mere debiasing in the sense that they force and block particular inferences in general, not just inferences based on our prior beliefs.

In step two, the subject uses *CS1* and the placeholder structure as the starting point for diverse *modeling processes*, including analogies (e.g., between gravitation and light), induction (e.g., if TWO is one more than ONE, and THREE is one more than TWO, then FOUR is one more than THREE), thought experiments, and limiting cases. These modeling processes are limited and directed by the constraints of the placeholder structure and its partial interpretation through *CS1*. The goal is to find and fit the placeholder structure to both the issue at hand and the constraints imposed by the potential partial interpretation through *CS1* and the placeholder structure. If they are successful, the cognizer can "fill in" the placeholder structure to generate a new and foreign *CS2* through these modeling processes. *CS2* can contain concepts foreign to *CS1* because the placeholder structure was not completely interpreted by *CS1*. Dutilh Novaes calls this modeling and interpretation, which transforms the placeholders into a symbol structure that applies to a new phenomenon and captures *CS2 re-semantification* (Dutilh Novaes 2012, p. 204).[15]

To sum up, bootstrapping is a process that can solve stubborn problems by introducing foreign concepts that were not accessible with the resources of the original conceptual system. The process is constrained by meaningless placeholder structures that can go beyond the original conceptual system but map onto the problem at hand and parts of the original system. Foreign concepts are introduced through modeling processes like analogies, induction or thought experiments about the placeholder structure, filling it with meaning. With this account of bootstrapping under our belt, we can now return to heavy-duty conceptual engineering.[16]

4 | HEAVY-DUTY CONCEPTUAL ENGINEERING VIA BOOTSTRAPPING

How would we go about conducting heavy-duty conceptual engineering and where do we start? The conceptual engineering literature often distinguishes between three stages of conceptual engineering: the *assessment* of a given (set of) concept(s), the *design* of one or several superior successor concept(s) and their *implementation*, establishing the use of the successor concept(s) in an epistemic community. In this section, we will adopt this tripartite structure and flesh out how each of these stages looks in the case of heavy-duty conceptual engineering.

We submit that heavy-duty conceptual engineering shares the assessment stage with other forms of conceptual engineering: it examines whether there are problems with the concepts currently in use. While Carey (2009, p. 414) limits these problems to (in-)consistency as "a major impetus for conceptual change" via bootstrapping, there is no reason to exclude other difficulties such as gaps in meaning, or a lack of fruitfulness or exactness (see Carnap, 1950). These and potentially other reasons can motivate episodes of heavy-duty conceptual engineering.

Whether the problem at hand can be solved via moderate conceptual engineering or requires heavy-duty conceptual engineering is determined by a procedure of trial and error, in which the conceptual engineer tests various stipulative or revisionary definitions for their propensity to solve

[15] See Nersessian (2008) for further details on modeling processes and how they figure in conceptual change.

[16] Rey (2014) has objected that bootstrapping must be circular because its success presupposes knowledge of the concepts one sets out to bootstrap. This objection overlooks the role of non-conceptual placeholder structures for bootstrapping. These placeholder structures provide bootstrappers with the constraints needed to develop foreign concepts without already being the bootstrapped concept (see Beck, 2017). A related worry is that bootstrapping cannot exclude deviant alternatives like the concept GRUE (Rey, 2014). However, as Beck (2017) notes, this worry is not specific to bootstrapping but applies globally to *any* inductive procedure.

the problem. If a candidate definition appears promising, it is further refined in the design stage; if no promising candidate definitions can be found, the problem may be stubborn and the design of foreign concepts is needed. If the problem is dismissed as misguided or meaningless, this is evidence that it is a stubborn problem. The fact that no solution can be found within the available conceptual system may lead some researchers to this conclusion. Thus, moderate and heavy-duty conceptual engineering will typically part ways at the end of the assessment stage.

At the design stage, heavy-duty conceptual engineering differs radically from light and moderate conceptual engineering. The design stage of the latter two involves the development of stipulative or revisionary definitions; the design stage of heavy-duty conceptual engineering involves heuristic bootstrapping. The complexity and heterogeneity of bootstrapping does not allow us to give a step-by-step IKEA-manual for conducting it. Nonetheless, we can draw on our account to give some general instructions for how to approach the design of collectively foreign concepts.

Insofar as heavy-duty conceptual engineering is a means of solving stubborn problems, the structure of these problems constrains the process of engineering concepts that can be used to solve them (Nersessian, 2008; Haueis & Slaby, 2022). The previous analysis has revealed a number of key elements for how this can be done. The first step will be to find a suitable placeholder structure. This structure may, but need not, resemble the structure of the symbols contained in *CS1*. The search for a placeholder structure is a crucial step in the overall process, because its internal structure as well as its partial interpretation guides its subsequent interpretation. A reasonable approach is likely to involve a period of trial and error, in which various placeholder structures that can be sourced from other domains with analogous problems are tested for their capacity to capture the known elements of the problematic phenomena. Often, even a well-chosen placeholder structure will need to be modified at several stages in the overall process, or several different placeholder structures will need to be integrated to address different aspects of the problematic phenomena.

Once a suitable placeholder structure has been articulated, it needs to be filled with content. This requires the use of modeling processes such as thought experiments, limiting cases, external analogies, or abductive reasoning. Analogies are helpful not only for discovering appropriate placeholder structures, but also for acquiring semantic knowledge of *CS2* – as opposed to the mere formal relations of its placeholders – by showing the similarities and differences to other known phenomena. For example, the analogy of interplanetary forces to light played an important role in Kepler's interpretation of the soul placeholder. Thought experiments help to draw out the implications of adopting a given analogous placeholder structure for one's target system *CS2*. These implications can then be assessed for their consistency with the constraints of the given problem, with one's general theoretical commitments, and with their mapping to existing data. If the process is successful, the bootstrapped *CS2* has the semantic expressive power to offer and express solutions to the problem at hand.

In real life cases, the processes described here are more complicated and chaotic. Frequently going back and forth between the stages is likely to be the norm – for example when a placeholder structure maps part of a problematic phenomenon but fails in other aspects. Moreover, owing to the immense diversity in topics, practical and theoretical problems, and antecedent knowledge, actual episodes of heavy-duty conceptual engineering will differ greatly from each other. Nonetheless, a better understanding of the nature of the task as well as its key components will facilitate the process – in particular, an awareness of the need for placeholder structures and modeling processes. For this reason, we recommend that would-be heavy-duty conceptual engineers also engage with relevant research in this area, e.g. on modeling in the process of conceptual change (Nersessian, 2008) or on the role of de- and re-semantification (Dutilh Novaes, 2012).

After successfully heuristically bootstrapping foreign concepts, the implementation of heavy-duty conceptual engineering involves *educative bootstrapping* to make the new concepts available to everybody in the community. Existing research shows that transitioning from an entrenched conceptual system to a foreign one is very hard and requires sophisticated methods of teaching and training. This is evidenced by how often science and math instruction fails to induce the conceptual shifts it aims for. Fortunately, there is a growing body of didactic and psychological research on what factors facilitate the process. For example, Smith (2007) develops and tests a curriculum for learning a discontinuous concept of matter. She explicitly models this process as a bootstrapping process from a *CS1* to a discontinuous *CS2*. Again, we recommend that those eager to implement the outcomes of heavy-duty conceptual engineering consult the relevant empirical literature (see also Moss & Case, 1999).

All in all, the three commonly identified stages of conceptual engineering – assessment, design, and implementation – have analogues in heavy-duty conceptual engineering. Whereas the assessment is more or less the same as in other types of conceptual engineering, the design and implementation stages require bootstrapping: heuristic bootstrapping in the design stage and educative bootstrapping in the implementation stage. In both cases, existing research helps to draw attention to factors that increase the likelihood of success.

5 | OBJECTIONS

We have argued that there is a type of particularly fundamental conceptual engineering, heavy-duty conceptual engineering, that has the capacity to expand the realm of thinkable thoughts and thereby contribute to solving stubborn problems. To further characterize this type of conceptual engineering and to distinguish it from other types of conceptual engineering, we have introduced the notions of foreign concepts and expressive power, arguing that heavy-duty conceptual engineering alters a community's semantic expressive power through the introduction of foreign concepts. Introducing foreign concepts calls for a process of concept introduction that goes beyond definitions. Inspired by Nersessian's and Carey's influential work, we have introduced the process of bootstrapping.

Our account of heavy-duty conceptual engineering commits us to the possibility of foreign concepts – concepts that exceed the semantic expressive power of the concepts we currently possess. In this final section, we will discuss objections against this theoretical commitment. This discussion will also serve to further clarify what heavy-duty conceptual engineering amounts to.

Two different objections against there being foreign concepts to engineer must be distinguished. The first is a *general* objection against the view that people *ever* acquire foreign concepts. This objection can be put as follows:

> *General objection*: For any person *S* and conceptual systems *CS1* and *CS2*, if *S* possesses *CS1* at time *t1* and *CS2* at time *t2*, then *CS2* contains only concepts that are either present in *CS1* or can be constructed by recombining the elements of *CS1*.

To deny that people ever acquire foreign concepts, and that they *a fortiori* cannot be engineered, is to embrace *radical nativism*: Fodor's infamous thesis that all concepts are either innate or constructed from innate concepts. This would imply that all conceptual engineering is either light or moderate. As many others have argued before us, radical nativism has absurd consequences: toddlers would already possess the conceptual resources to build concepts like SPACETIME, COM-

PUTER VIRUS, or TRANSFINITE NUMBER. We reject this "mad dog nativism" (Rey, 2014). Once radical nativism is out of the way, however, one must accept that human beings *sometimes* acquire foreign concepts, that these can *in principle* be engineered, and hence that the *General objection* is misguided.

The second objection targets heavy-duty conceptual engineering more specifically:

> *Specific objection*: For any person *S* who possesses *CS1* at time *t1*: if it is possible for *S* to engineer a concept *c* via heavy-duty conceptual engineering, then *c* can be constructed via recombinations of the elements of *CS1*.

This objection can be bolstered by pointing to the Ramsey-Carnap-Lewis method (Ramsification) of defining theoretical terms using only one's old *CS1* by stating the role that these terms play within a given theory (Lewis, 1970).

Ramsification is a method to define the terms of an *already developed* theory. Once we have the theory, its theoretical terms might be translated into *CS1* through a so-called Ramsey sentence that represents the theory. It is, however, not clear that Ramsification from a *CS1* is always possible: Carnap famously failed to define the predicate "is at" (Quine, 1951) or dispositions like "solubility" (Leitgeb & Carus, 2023) within his conceptual system. Not all theories are expressible in all conceptual systems, even if they are highly sophisticated.

Even granting that there is a Ramsey-sentence capturing concept *c* in *CS1*, this is of little help to would-be heavy-duty conceptual engineers. A heavy-duty conceptual engineer does not yet have a complete theory that specifies the role of the concept that she aims to define and that she needs to be able to capture the term and formulate the Ramsey sentence. While a Ramsey sentence that characterizes the target concept may exist *in abstracto*, she first needs to bootstrap the concept and its associated theory to be able to find it.

We take this to be one of the hallmark features of heavy-duty conceptual engineering. Moderate conceptual engineering defines concepts that serve specific theoretical roles within already developed theories. By contrast, heavy-duty conceptual engineers need to engage in bootstrapping processes in which theory and concepts are developed in tandem. Neither does the aspiring heavy-duty conceptual engineer possess the concepts she needs to find the theory expressed by the Ramsey sentence; nor does she know the Ramsey sentence she would need to define the concepts the Ramsey sentence contains by Ramsification.

Take *S*'s *CS1* at time *t1*, a theory *T* that features the theoretical term *t*, and *CS2* that is the union of *CS1* and the concept expressed by *t*. Then the following holds: even if ideally rational, *S* can find the Ramsey sentence expressing *T* to Ramsify *t* in terms of *CS1* at time *t1* only if *S* knows *T* at *t1* (or can otherwise make use of *T*). This is just to say that, in order to interpret t via Ramsification, *S* needs to be able to make use of the theoretical role that *T* assigns to *t*, i.e. know and possess the associated concept. This condition is not satisfied when it comes to heavy-duty conceptual engineering. We do not already know a fully formed theory or Ramsey sentence when we develop foreign concepts. Note that the point we are making here is not merely psychological. Even an ideally rational agent will not be able to define the concept expressed by *t* without first knowing the prerequisite theory *T*.[17]

[17] Of course, this is not to deny that an empirically omniscient agent, who would know *T*, could use Ramsification to define *t*. This method is not for omniscient agents, who already possess all concepts, but for beings like us.

6 CONCLUDING REMARKS

This paper explores a radical idea: via conceptual engineering, we can gain access to previously unthinkable thoughts. We call this heavy-duty conceptual engineering. In the introduction we asked a row of questions: What can we hope to achieve by doing conceptual engineering? What does it mean for a thought to be "unthinkable"? How can previously unthinkable thoughts "become thinkable"? What does heavy-duty conceptual engineering actually amount to? Is heavy-duty conceptual engineering possible in practice? And if so, how can it be put in the service of tackling stubborn problems?

In a nutshell, our answers are as follows. Heavy-duty conceptual engineering is a kind of conceptual engineering that alters an epistemic community's semantic expressive power by making formerly unthinkable thoughts thinkable. A thought is unthinkable if it is constituted by at least one foreign concept. A foreign concept is a concept that is neither available nor latent. Foreign concepts can be made available through bootstrapping. Heavy-duty conceptual engineering uses heuristic bootstrapping at its design-stage and educative bootstrapping at its implementation stage. Numerous detailed case studies of conceptual change in the history of science establish the possibility of heuristic bootstrapping. Moreover, there is a body of research in cognitive science and didactics on educative bootstrapping. This research can be used to attain a more concrete understanding of how heavy-duty conceptual engineering can help us to introduce the foreign concepts needed to develop solutions to stubborn problems.

Where does this leave us with respect to the bold promise that conceptual engineering has the power to change what is thinkable for subjects like us? Is the optimism formulated by Burgess and Plunkett (2013) justified after all? Or is Eklund (2021) right that their rhetoric is exaggerated and ought to be tempered? The account of heavy-duty conceptual engineering we presented in this paper provides a clear picture of what is at stake in this disagreement. It has emerged that Burgess & Plunkett's rhetoric can indeed be justified because heavy-duty conceptual engineering is actually possible, both in theory and in practice. Nonetheless, Eklund's plea for sobriety is not unreasonable either. Our discussion has shown just how difficult heavy-duty conceptual engineering is. Moreover, there is a principled reason to believe that many of the instances of conceptual engineering that we find in philosophy papers are not of the heavy-duty kind: They are presented in the form of definitions – the hallmark of moderate and light conceptual engineering.

Our discussion of heavy-duty conceptual engineering allows for several further conclusions that future debates should take into account. First of all, it emerges that not all conceptual engineering is created equal. There are stark differences between light, moderate and heavy-duty conceptual engineering, neglecting which leads to distorted views of the goals and the processes involved in conceptual engineering. On the one hand, not all conceptual engineering has the power to expand the realm of the thinkable; indeed, many of its instances – those we group under *light* and *moderate* conceptual engineering – do not. On the other hand, when critics such as Max Deutsch remark that conceptual engineering is simply about "replac[ing] longer descriptions [...] with a shorter, single term" (Deutsch, 2020, p. 3945), it should now be clear that this applies at best to light conceptual engineering.[18]

Second, properly acknowledging heavy-duty conceptual engineering as a distinctive method to confront stubborn problems enhances the toolkit of would-be conceptual engineers. Whereas light and moderate conceptual engineering proceed mostly by revisionary definitions, heavy-duty

[18] But see (Koch, 2021) for criticism also of this claim.

conceptual engineering shows us how to proceed if these methods do not suffice to solve the problem at hand.

Finally, we submit that the study of heavy-duty conceptual engineering in philosophy provides a valuable but neglected area of research in its own right. While philosophers of science have made various attempts to identify and describe instances of heavy-duty conceptual engineering in science, there has been almost no research on what role it plays in philosophy. A more in-depth study of how heavy-duty conceptual engineering *has been* applied in philosophy will shed further light on how the method *can* or *should be* applied to future cases. We take this to be an important new research program.[19]

ACKNOWLEDGMENTS

This paper has long been in the making, and we are grateful for all the feedback we have received along the way. We thank Matti Eklund, Philipp Haueis, Alex Kocurek, and Christian Nimtz as well as three anonymous referees for *Noûs* for their helpful comments on earlier versions of this paper. We also thank Thomas Grundmann for lending us the office in which the gist of this paper was developed. Additionally, we thank the audiences of Bielefeld's philosophy research colloquium (2021), the EXTRA.8 research colloquium in Bochum (2022), the MüBiOs (online) workshop (2022), the Conceptual Engineering Online Seminar (2022), the EXTRA.3 workshop on conceptual engineering and its place in philosophical methodology (2022), GAP.11 (2022), and the workshop on philosophical and psychological perspectives on conceptual change at UNAM (2023).

ORCID

Steffen Koch https://orcid.org/0000-0002-9360-799X
Jakob Ohlhorst https://orcid.org/0000-0001-8630-6358

BIBLIOGRAPHY

Andersen, H., Barker, P., & Chen, X. (2006). *The Cognitive Structure of Scientific Revolutions*. Cambridge: Cambridge University Press.

Beck, J. (2017). Can bootstrapping explain concept learning? *Cognition*, *158*, 110–121. https://doi.org/10.1016/j.cognition.2016.10.017

Berkovitz, J. (2016). Action at a Distance in Quantum Mechanics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016). https://plato.stanford.edu/archives/spr2016/entries/qm-action-distance/

Boden, M. A. (1990). *The Creative Mind: Myths and Mechanisms*. New York: Basic Books.

Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, *81*, 1211–1241. https://doi.org/10.1007/s10670-015-9791-5

Burgess, A., & Plunkett, D. (2013). Conceptual Ethics I. *Philosophy Compass*, *8*, 1091–1101. https://doi.org/10.1111/phc3.12086

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

Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

[19] As a conjecture about what might be plausible instances of heavy-duty conceptual engineering in the more recent history of philosophy, we suggest studying the development of modal logic and the proposed theory of integrated information theory (Tononi, 2004). Both are often viewed as solutions to problems whose solutions are not expressible within our original conceptual frameworks. Additionally, both propose purely formal placeholder structures that are subsequently interpreted – set theory in the first case and complex mathematical formulae in the second. This being said, convincingly arguing that either of these cases fits the structure of heavy-duty conceptual engineering would require in-depth reconstructions analogous to the efforts of Gentner (2002) or Nersessian (2008).

Carey, S. (2014). On Learning New Primitives in the Language of Thought: Reply to Rey. *Mind & Language*, *29*, 133–166. https://doi.org/10.1111/mila.12045

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Chalmers, D. J. (2020). What is conceptual engineering and what should it be? *Inquiry*, 1–18. https://doi.org/10.1080/0020174X.2020.1817141

Chalmers, D. J. (2022). *Reality+. Virtual worlds and the problems of philosophy*. London: Allen Lane.

Christensen, D. (2004). *Putting logic in its place*. Oxford: Oxford University Press.

Deutsch, M. (2020). Speaker's reference, stipulation, and a dilemma for conceptual engineers. *Philosophical Studies*, *177*, 3935–3957. https://doi.org/10.1007/s11098-020-01416-z

Dutilh Novaes, C. (2012). *Formal Languages in Logic*. Cambridge: Cambridge University Press.

Egré, P., & O'Madagain, C. (2019). Concept utility. *The Journal of Philosophy*, *116*, 525–554. https://doi.org/10.5840/jphil20191161034

Eklund, M. (2021). Conceptual Engineering in Philosophy, in J. Khoo & R. Sterken (Eds.), *The Routledge Handbook of Social and Political Philosophy of Language* (pp. 15–30). Routledge.

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. New York: Oxford University Press.

Gentner, D. (2002). Analogy in scientific discovery: The case of Johannes Kepler. In L. Magnani & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 21–39). New York: Kluwer Academic/Plenum.

Haslanger, S. (2012). *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.

Haueis, P. & Slaby, J. (2022). The humanities as conceptual practices: The formation and development of high-impact concepts in philosophy and beyond. *Metaphilosophy*, *53*, 385–403. https://doi.org/10.1111/meta.12551

Isaac, M. G. (2023). Which concept of concept for conceptual engineering? *Erkenntnis*, *88*, 2145–2169. https://doi.org/10.1007/s10670-021-00447-0

Isaac, M. G., Koch, S., & Nefdt, R. (2022). Conceptual engineering: a road map to practice. *Philosophy Compass*, *17*, e12879. https://doi.org/10.1111/phc3.12879

Koch, S. (2021). There is no dilemma for conceptual engineering. Reply to Max Deutsch. *Philosophical Studies*, *178*, 2279–2291. https://doi.org/10.1007/s11098-020-01546-4

Koch, S., Löhr, G., & Pinder, M. (2023). Recent work in the theory of conceptual engineering. *Analysis*, *83*, 589–603. https://doi.org/10.1093/analys/anad032

Koch, S., & Lupyan, G. (2024). What is conceptual engineering good for? The argument from nameability. *Philosophy and Phenomenological Research*, 1–18. https://doi.org/10.1111/phpr.13110

Kuhn, T. (1996). *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.

Leitgeb, H. & Carus A. (2023). Rudolf Carnap. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition). https://plato.stanford.edu/archives/sum2023/entries/carnap/

Lewis, D. (1970). How to Define Theoretical Terms. *Journal of Philosophy*, *67*(13), 427–446. https://doi.org/10.2307/2023861

Margolis, E., & Laurence, S. (2019). Concepts. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition). https://plato.stanford.edu/archives/spr2021/entries/concepts/

Moss, J., & Case, R. (1999). Developing children's understanding of rational numbers: a new model and an experimental curriculum. *Journal for Research in Mathematics Education*, *30*, 122–147. https://doi.org/10.2307/749607

Nado, J. (2023). Classification procedures as the targets of conceptual engineering. *Philosophy and Phenomenological Research*, *106*, 136–156. https://doi.org/10.1111/phpr.12843

Nersessian, N. (2008). *Creating Scientific Concepts*. Cambridge, MA: MIT Press.

Pinder, M. (2020). What ought a fruitful explicatum to be? *Erkenntnis*, *87*, 913–932. https://doi.org/10.1007/s10670-020-00223-6

Pinder, M. (2021). Conceptual engineering, metasemantic externalism and speaker-meaning. *Mind*, *130*, 141–163. https://doi.org/10.1093/mind/fzz069

Queloz, M. (2022). Function-Based Conceptual Engineering and the Authority Problem. *Mind*, *131*, 1247–1278. https://doi.org/10.1093/mind/fzac028

Quine, W. V. O. (1951). Two Dogmas of Empiricism. *The Philosophical Review*, *60*, 20–43. https://doi.org/10.2307/2181906

Rey, G. (1983). Concepts and stereotypes. *Cognition*, *15*, 237–262. https://doi.org/10.1016/0010-0277(83)90044-6

Rey, G. (2014). Innate and learned: Carey, mad dog nativism, and the poverty of stimuli and analogies yet again. *Mind & Language*, *29*, 109–32. https://doi.org/10.1016/0010-0277(83)90044-6

Ritchie, K. (forthcoming). Labeling unlabeled identities. In M.G. Isaac, S. Koch, & K. Scharp (Eds.), *New Perspectives on Conceptual Engineering Vol. 3: Applied Conceptual Engineering*. Dordrecht: Springer.

Scharp, K. (2020). Philosophy as the study of inconsistent concepts. In A. Burgess, H. Cappelen, & D. Plunkett (eds.), *Conceptual Engineering and Conceptual Ethics* (pp. 396–416). Oxford: Oxford University Press.

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

Simion, M. & Kelp, C. (2020). Conceptual innovation, function first. *Nous*, *54*, 985–1002. https://doi.org/10.1111/nous.12302

Smith, C. L. (2007). Bootstrapping Processes in the Development of Students' Commonsense Matter Theories: Using Analogical Mappings, Thought Experiments, and Learning to Measure to Promote Conceptual Restructuring. *Cognition and Instruction*, *25*, 337–398. https://doi.org/10.1080/07370000701632363

Smithies, D. (2015). Ideal rationality and logical omniscience. *Synthese*, *192*, 2769–2793. https://doi.org/10.1007/s11229-015-0735-z

Sterken, R. K. (2020). Linguistic Intervention and Transformative Communicative Disruptions. In A. Burgess, H. Cappelen, & D. Plunkett (Ed.), *Conceptual Engineering and Conceptual Ethics* (pp. 417–434). Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198801856.003.0020

Thagard, P. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neurosci*, *5*, 42. https://doi.org/10.1186/1471-2202-5-42

How to cite this article: Koch, S., & Ohlhorst, J. (2024). Heavy-duty conceptual engineering. *Noûs*, 1–19. https://doi.org/10.1111/nous.12538