



Engineering what? On concepts in conceptual engineering

Steffen Koch¹ 

Received: 25 July 2020 / Accepted: 9 September 2020 / Published online: 18 September 2020
© The Author(s) 2020

Abstract

Conceptual engineers aim to revise rather than describe our concepts. But what are concepts? And how does one engineer them? Answering these questions is of central importance for implementing and theorizing about conceptual engineering. This paper discusses and criticizes two influential views of this issue: *semanticism*, according to which conceptual engineers aim to change linguistic meanings, and *psychologism*, according to which conceptual engineers aim to change psychological structures. I argue that neither of these accounts can give us the full story. Instead, I propose and defend the *Dual Content View of Conceptual Engineering*. On this view, conceptual engineering targets concepts, where concepts are understood as having two (interrelated) kinds of contents: referential content and cognitive content. I show that this view is independently plausible and that it gives us a comprehensive account of conceptual engineering that helps to make progress on some of the most difficult problems surrounding conceptual engineering.

Keywords Conceptual engineering · Semantic externalism · Concepts · Dual content · Reference · Philosophical and psychological theories of concepts

1 Introduction

Conceptual engineers think that philosophy should involve the critique and improvement of the concepts we use—both within and outside of philosophy. The potential impacts of this approach are interesting and far-reaching. Few things in the world are more important to human cognition and interaction than our concepts. They shape how we think about the world, how we communicate with each other, how we pursue our personal lives, and how we organize our society. Philosophy understood as conceptual engineering clearly has a lot of potential.

✉ Steffen Koch
steffen.koch@rub.de

¹ Institut für Philosophie II, Ruhr-Universität Bochum, Universitätsstraße 150, 44801 Bochum, Germany

But what is the relevant notion of concepts? And what could it possibly mean to engineer them? These questions are of central importance to both the theory and the practice of conceptual engineering; but, as I'll argue in this paper, they haven't been answered satisfactorily.¹ There are currently two main approaches to these questions in the literature. According to what I'll call *semanticism*, conceptual engineering is concerned with word meanings. On this view, to do conceptual engineering is to advocate and implement changes in what our words mean (Cappelen 2018; Thomasson 2020). By contrast, according to what I'll call *psychologism*, conceptual engineering is concerned with the psychological structures that explain our mental and linguistic behavior. On this view, to do conceptual engineering is to advocate and implement changes in how people classify things, what inference patterns they are drawn to, and under what circumstances they use particular linguistic expressions (Machery 2017; Isaac 2020).

In this paper, I argue that each of these approaches is, at best, only part of the story. Conceptual engineering is typically described as a means of achieving at least two goals at once: the *semantic goal* of changing what certain terms or expressions mean in a language, and the *practical goal* of effecting certain real-world changes, such as changes in linguistic practices or people's classificatory behavior. But neither semanticism nor psychologism provides an adequate means of achieving both of these goals at once. Therefore, if we construe conceptual engineering along the lines of semanticism or psychologism, it cannot live up to the expectations of its practitioners. Moreover, semanticism and psychologism both run the risk of detaching semantic facts from linguistic practices, which can lead to confusion and miscommunication. To avoid these unwanted consequences, and to turn conceptual engineering into a means of achieving both the semantic and the practical goal, a hybrid view is needed. The principal aim of this paper is to motivate, develop, and defend such a view.²

The basic idea behind this view is that conceptual engineering is concerned with concepts, which I take to be mental representations with dual contents: referential content on the one hand, and cognitive content on the other.³ Within this framework, there are two dimensions along which concepts can be defective and improved: the dimension of *what* they represent, e.g. classes of people, objects, events, etc., and the dimension of *how* they represent the members of their extension. I will argue that by viewing conceptual engineering as operating on concepts with dual contents, we can accommodate both the semantic and the practical goal at once. In order to achieve the semantic goal, we ought to engineer referential contents; in order to achieve the practical goal, we ought to engineer cognitive contents. This framework thus allows

¹ There is a wide and a narrow sense of 'conceptual engineering' used in the literature. In the wide sense, conceptual engineering is the general practice of revising, replacing, introducing or eliminating concepts; in the narrow sense, it is limited to revising certain already existing concepts. Examples include Haslanger (2000), Scharp (2013), and Fassio and McKenna (2015). This paper focuses on conceptual engineering in the narrow sense.

² I am not the first author to propose a dual representation view of conceptual engineering. Most notably, Sarah Sawyer has done so in a series of recent papers (2018, 2020a, 2020b). However, the motivation for and the details of her view are rather different from mine (see Sect. 6).

³ See Margolis and Laurence (1999), Carey (2009), Weiskopf (2009), and Del Pinal (2015) for contemporary articulations of dual content views outside the conceptual engineering debate. See Evans (1973) and Devitt (1981) for early predecessors of the view.

us to remedy the main difficulties of semanticism and psychologism. What is more, it allows us to make progress on some of the most contested issues in the recent literature, such as whether conceptual engineering is within our control (Cappelen 2018), whether there can be amelioration as revelation (Haslanger 2006), and how to respond to Strawson’s challenge (Strawson 1991/1963; Prinzing 2017; Cappelen 2018; Sawyer 2018; Nado 2019).

Here is the roadmap. In Sect. 2, I argue that conceptual engineering should be viewed as having both semantic and practical goals, and that existing views fail to capture them. Then, in Sects. 3, 4, 5, I present my own positive view. Section 3 starts by explaining and motivating the idea of dual contents. Section 4 lays out my version of the dual content view of concepts and shows how it can be transferred to linguistic meanings. Section 5 demonstrates what engineering concepts looks like on such a view. Finally, Sect. 6 points out the advantages that we can expect from holding the view outlined in this paper.

2 Conceptual engineering: major goals and approaches

In this section, I will argue for the following three claims. First, conceptual engineering ought to be construed as an activity with the two tandem goals of changing semantic values and changing people’s usage of words and concepts. Second, given more or less any plausible metasemantic theory, semantic values and usage can come apart. Finally, for this very reason, the two most popular general approaches to conceptual engineering—semanticism and psychologism—both fail to provide adequate frameworks for theorizing about and practicing conceptual engineering. The subsequent sections take up the challenge of constructing a more comprehensive alternative.

Conceptual engineering is often described as the activity (or a method) of changing what certain words or expressions mean in a language. For instance, Cappelen (2018), who takes meanings to consist of intensions and extensions, argues that conceptual engineering “should be seen as having as its goal to change extensions and intensions of expressions” (p. 61). Haslanger (2020) describes her strategy concerning ‘race’ and ‘gender’ as “appropriat[ing] terminology” and improving “our conceptual and linguistic tools” (p. 8). And, using the example of ‘truth’, Eklund (2015) describes conceptual engineering “as a study of what concept best plays the theoretical role of our concept of truth and what features this concept has” (p. 376). On the received view, concepts are word meanings; so what Eklund describes in this passage can be viewed as the study of which meanings we should ascribe to the words we use, given the theoretical purposes at hand. If our current meanings turn out to be objectionable, then we should try to change them. This approach to conceptual engineering also surfaces in many concrete conceptual engineering proposals, such as Manne’s proposal to change the meaning of ‘misogyny’ (Manne 2018). The list could be continued. It therefore seems reasonable to construe conceptual engineering as an activity with the following goal:

Semantic goal: Conceptual engineers (at least sometimes) aim to change the meanings of words and thereby change the truth values of sentences (e.g. change

sentences like ‘Jackson is Lisa’s father’ or ‘Grace is a woman’ from being false to being true, or *vice versa*).

However, conceptual engineers also point out that they take their endeavors to have significant practical effects beyond semantics. In this vein, Cappelen (2018) argues that conceptual engineering will (and should) not merely change the semantic features of our representational devices, but also have first-order effects on the world we live in (pp. 137ff.). Similarly, Burgess and Plunkett (2013) make clear that conceptual ethics has an impact “not only [on] 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 will therefore constrain “what we can hope to accomplish in the world” (pp. 1096–1097). Likewise, Haslanger (2020) argues that social emancipation, to which she intends her revisionary analyses of race and gender to contribute, “must be a collective effort and change more than minds” (p. 7).

It is implausible that conceptual engineering could have such a significant practical impact merely by accomplishing its semantic goal. For people to entertain different thoughts, form different desires, and act differently, it is crucial that they actually change the way they think about certain categories and how they mentally and linguistically classify objects, people, events, etc. This practical attitude is reflected in many writings on the topic. Sawyer (2020b), for example, describes conceptual engineering as “a form of theorizing that involves a proposed change in linguistic practice” (p. 2), and in Manne’s (2018) proposal for engineering ‘misogyny’, considerations about the effects of calling an agent ‘misogynist’ play an important role in her argument for her proposed view. All this suggests that we should also ascribe the following goal to conceptual engineers:

Practical goal: Conceptual engineers (sometimes) aim to change how people think about objects, how they classify them, and how they use words (e.g. by getting people to stop calling whales ‘fish’, or start calling certain acts ‘misogynistic’).

If these two goals are indeed the goals of conceptual engineers, then how are they connected? First of all, it would seem that each of these goals is important in its own right. The semantic goal is important because it often matters what meaning a word has, and accordingly which sentences come out as true in certain circumstances. The practical goal is important because conceptual engineering cannot unleash its full potential unless it has real-world effects. Merely changing semantic values will often not be enough; what is needed is a change in how people actually think about and classify things and how they use the terms in question.

But there is another, perhaps even stronger reason for conceptual engineers to pursue both goals at once. The practical goal and the semantic goal are *tandem goals*: if one aims to achieve one of them, then one cannot neglect the other without risking unfavorable consequences. If we strive for the semantic goal, but neglect the practical goal, then we promote *linguistic obliviousness*: we change the meanings of words used by people without many of them noticing and changing their linguistic practice accordingly. In contrast, if we strive for the practical goal, but neglect the semantic goal, then we promote *linguistic confusion*: we lead people to use words incorrectly

and thus to assert many falsehoods. Conceptual engineers should avoid promoting both *linguistic obliviousness* and *linguistic confusion*. This is a strong reason to aim for both goals simultaneously, and not to neglect one of them in favor of the other.⁴

To see this more clearly, consider an example. Suppose Amy, a marine biologist, strongly desires to change the meaning of ‘fish’. As a matter of fact, in Amy’s linguistic community, ‘fish’ has whales in its extension, and everybody happily applies this term to whales. But Amy thinks that, all things considered, it would be better if whales were not in the extension of ‘fish’. Therefore, she strives for the semantic goal of changing the meaning of ‘fish’. Suppose that she is successful. If she doesn’t also work toward changing how people actually apply the term ‘fish’, i.e. the practical goal, she will have promoted linguistic obliviousness. Many formerly competent users of the term ‘fish’ no longer know what they are talking about: they will think they are talking about whales whereas in fact they aren’t. The same holds, *mutatis mutandis*, if Amy strives for the practical goal while neglecting the semantic goal.

So far so good; but why think that the two goals can come apart in such unfavorable ways? Doesn’t achieving the semantic goal imply achieving the practical goal, and *vice versa*? Not quite. To see this, consider the following principle:

Divergence: What a term t actually means in a language L can diverge from what many (or even most) speakers of L think t means and from their dispositions to use t .

Divergence follows from semantic externalism. On Kripke’s (Kripke 1970/1980) causal-historical view of reference, a name or a natural kind term refers to the individual or kind to which it was introduced. Once reference is fixed in this way, it is then sustained by chains of communication stretching from the term’s introduction to its current use—even if speakers have divergent beliefs and dispositions concerning that term. Similarly, according to Putnam’s (1975) version of externalism, depending on the environment they happen to live in, two people with the same beliefs and dispositions regarding e.g. ‘water’ may nevertheless refer to different entities. And, according to Burge’s (1979) social externalism, lay people and experts with radically different beliefs and dispositions regarding certain terms may nevertheless refer to the same entities when they utter those terms. All these and other versions of semantic externalism imply *Divergence*.⁵ And many internalist views will give rise to some version of *Divergence* as well. Consider e.g. views on which meanings are determined by reflective or idealized usage (Sawyer 2020b). Although not entirely uncontroversial, there are strong reasons for assuming that some version of *Divergence* holds.⁶

⁴ Note that I am not here assuming that conceptual engineering is truth-directed. What I assume is merely that conceptual engineers should shy away from *worsening* the proportion of true assertions within a community. This assumption is also bolstered by considerations about the normativity of meaning (see Glüer and Wikforss 2009 for a useful overview).

⁵ The strength of *Divergence*, i.e. how many speakers of L may have diverging beliefs and dispositions, and how much their beliefs and dispositions may diverge from the actual references, will depend on the details of the view. Cappelen (2018) endorses an extremely strong version of *Divergence*, according to which even all speakers of a language can have divergent beliefs, dispositions and usage (p. 63). Other externalist views are less extreme, but all of them entail some version of *Divergence*.

⁶ One might object against the above line of reasoning that e.g. Kripke and Putnam endorse *Divergence* only for a limited number of expressions such as proper names or natural kind terms. However, notice that

Given *Divergence*, the meaning of a word is one thing; how people generally use that word, or how they classify things mentally, is another. So achieving the semantic goal does not imply achieving the practical goal, nor *vice versa*. But given that both of these goals are important in their own right, and furthermore that we risk promoting linguistic obliviousness or linguistic confusion if we strive for one of them while neglecting the other, it is an important and non-trivial challenge to construe conceptual engineering in such a way that it can incorporate, and be a means of achieving, both goals. In the remainder of this section, I will argue that the two most popular general approaches to conceptual engineering fail to meet this challenge.

The literature about conceptual engineering is roughly split in two camps. Philosophers in the first camp advocate instances of what we might call *semanticism*:

Semanticism: To engage in conceptual engineering is to advocate and implement changes in what our words mean.

Notice that this formulation of semanticism is neutral about what meanings are. The most popular version of semanticism is Cappelen's (2018) *Austerity Framework*, according to which conceptual engineering targets the intensions and extensions of selected expressions of a language. But for present purposes, it doesn't really matter whether meanings are intensions, extensions, references, definition-like, or whatever else.

If to engage in conceptual engineering is to advocate and implement changes in what our words mean, then conceptual engineering is a good means of achieving the semantic goal. But semanticism does poorly with respect to the practical goal. Given that semanticism targets meanings rather than mental or linguistic behavior, and that, as per *Divergence*, the two can come apart to a large extent, there is no guarantee that by changing the semantic facts, we also change people's mental and linguistic classifications. So it is questionable whether semanticism construes conceptual engineering in a way that makes it an adequate means of achieving the practical goal. As the practical goal is important in its own right, this tells us that semanticism doesn't give us the full story. Furthermore, given the danger of promoting linguistic obliviousness by striving for the semantic goal while neglecting the practical goal, semanticism turns conceptual engineering into a practice of dubious value.⁷

Philosophers in the other major camp advocate instances of what we might call *psychologism*:

Footnote 6 continued

Divergence is not tied to any of the specific views of reference determination mentioned in the main text. Moreover, there is increasing pressure to extend externalism to terms other than proper names and natural kind terms, including, for example, social kind terms such as 'woman' or 'race' (cf. Putnam 1975, p. 242; Burge 1986; Haslanger 2006; Löhr 2019; Sawyer 2020b).

⁷ It is worth mentioning that Cappelen (2018), one of the leading defenders of semanticism, also discusses lexical effects, which he takes to be the non-semantic and non-pragmatic effects of words (p. 123). Can lexical effects play the role of the psychological component needed to accommodate the practical goal? There are two reasons for doubt. First, lexical effects are not plausibly construed as the mental content that grounds our mental and linguistic classifications. If anything, lexical effects are explained by this mental content. Second, Cappelen does not take lexical effects to be something that conceptual engineers do or should ameliorate.

Psychologism: To engage in conceptual engineering is to advocate and implement changes in how people use words, classify objects around them, and draw inferences about those objects.

Again, this formulation of the view is fairly neutral, e.g. with respect to what grounds our linguistic behavior, what mental categorizations are, what concepts are, etc. The most popular version of psychologism is Machery's (2017) view, according to which conceptual engineering targets concepts, where concepts are understood as sets of beliefs (or belief-like states) that are retrievable by default from long-term memory. On Machery's view, concepts ground our classificatory behavior. If we want to change e.g. how people use words, or how they classify things mentally, then we ought to change their concepts. The details of Machery's view aren't essential to the definition of psychologism, however.

The problem with psychologism parallels the one with semanticism: if to engage in conceptual engineering is to advocate and implement changes in how people use words and how they mentally classify the objects around them, e.g. by changing their (psychologically construed) concepts, then conceptual engineering is a good means of achieving the practical goal. But given that psychologism targets psychological structures rather than word meanings, and that, as per *Divergence*, the two can come apart, there is no guarantee that changing the psychological structures of even large parts of a linguistic community will change the semantic facts. It is therefore questionable whether psychologism construes conceptual engineering in a way that makes it an adequate means of achieving the semantic goal. But as the semantic goal is important in its own right, this tells us that psychologism doesn't give us the full story either. Furthermore, in light of the danger of promoting linguistic confusion by striving for the practical goal while neglecting the semantic goal, psychologism turns conceptual engineering into a practice of dubious value.^{8,9}

This concludes the argument of this section. Let me stress that it is not meant to be a knockdown argument against semanticism or psychologism. Nonetheless, the above considerations lend *prima facie* plausibility to the idea that conceptual engineering should be viewed as striving simultaneously for the semantic and the practical goal. And if this is right, it gives us reason to seek an account of conceptual engineering on which it is a good means of achieving them both. Developing such an account is the task that I turn to now.

⁸ Notice that my claim is compatible with the idea that, depending on the specifics of the case, working toward the practical goal might sometimes be a reasonable means of working toward the semantic goal as well. But given that there is no guarantee of this, neither one of them is reducible to the other.

⁹ At this point, I should mention that there is a third view besides semanticism and psychologism. Pinder (2019) argues that the goal of conceptual engineering is merely to change what (a certain group of) people speaker-mean by certain expressions, rather than to change semantic meanings or concepts. Since the connection between speaker-meaning and semantic meaning is rather loose, however, this view falls prey to a similar objection as the one I raised against psychologism.

3 Motivating dual contents

We have seen that neither psychologism nor semanticism about conceptual engineering can give us the full story. The goal of the following sections is to provide a more comprehensive metasemantic framework for conceptual engineering that can. The core idea of this framework is that, for the purposes of conceptual engineering, concepts should be viewed as having two kinds of content—referential content (r-content) on the one hand, and cognitive content (c-content) on the other. Within this framework, conceptual engineering can be described as operating (primarily) on concepts; but since there are two kinds of conceptual content, there are also two ways of engineering concepts. Taken together, these two ways of engineering concepts allow us to view conceptual engineering as a means of achieving both the semantic and the practical goal.

Before I get to develop this view in further detail, let me address one caveat. Semanticists who agree with the arguments in the last section might be tempted to resort to the distinction between concepts and conceptions to adopt a view similar to the one I will defend. A conception of x is usually understood as the set of beliefs we have about x . Thus, semanticists might argue that whereas achieving the semantic goal requires engineering concepts, achieving the practical goal requires engineering conceptions. In order to avoid linguistic confusion and obliviousness, conceptual engineers ought to keep concepts and conceptions in sync. Let us call this view *semanticism-cum-conceptions*.

A version of semanticism-cum-conceptions has recently been defended by Sawyer (2018, 2020a, b). I will discuss the specifics of her subtle account in Sect. 6. For now, suffice it to say the following. Semanticism-cum-conceptions is not semanticism, so my arguments against semanticism are unaffected by whether or not semanticism-cum-conceptions can be rendered plausible. Whether it is indeed a plausible alternative to the view I am about to develop depends on how its details are spelled out. Since there is no uniform way of doing so, it is difficult to give a general assessment of the view at this point. I will therefore return to this issue in Sect. 6 and set it aside for now.

In the remainder of this section, I will introduce the two major sources that my view draws on—hybrid theories of reference on the one hand, and dual content accounts of concepts on the other. I will argue that both of these two strands of metasemantic theorizing provide further support for the dual content view. Lastly, I will discuss whether and in what respects the two contents alluded to here should be viewed as conceptual contents.

3.1 Hybrid theories of reference

In *Naming and Necessity*, Kripke articulates the outline of what has become known as the causal theory of reference about proper names and natural kind terms.¹⁰ As I will mostly be concerned with kind terms here, I'll formulate the view with respect to them:

¹⁰ To be fair, Kripke himself speaks of a “picture” rather than “theory”.

Reference-Fixing. The reference of a kind term t is established via ostension or a reference-fixing description of a set of paradigm instances of a kind k . A mechanism of implicit or explicit generalization makes sure that all and only tokens of k are in the extension of t . Whether or not a token belongs to k is a matter of shared essential properties.

Reference-Preservation. Once reference is established, it can easily be transmitted via chains of communication. For a speaker to gain the ability to refer to a kind k with the term t , it is more or less sufficient that she hears someone use t correctly, and that she intends to use t just as this person did.

The main rival to Kripke's view is descriptivism. Roughly, according to descriptivism, the reference of a term t as used by a speaker S is determined by the descriptive content that S associates with t . By contrast, on Kripke's view, what a speaker knows or associates with a term is irrelevant for what this term refers to. Speakers can be in massive error about what their terms refer to, but still succeed in using them. Many philosophers agree that Kripke's arguments against descriptivism, i.e. the modal argument and the argument from ignorance and error, are successful, and that his own view offers an attractive alternative. However, even among those who sympathize with Kripke's main ideas, some have raised doubts about whether associated content on the speaker's side can be entirely dispensed with. In particular, Gareth Evans (1973) and Michael Devitt (1981) noted early on that Kripke's view predicts reference to be constant in cases where intuitively it has switched.¹¹

Evans and Devitt have both formulated solutions to this problem that are very similar in spirit. As Evans (1973) diagnoses, the problem is that Kripke "has mislocated the causal relation; the important causal relation lies between [an] item's states and doings and the speaker's body of information—not between the item's being dubbed with a name and the speaker's contemporary use of it" (p. 197). On an Evansian account, the reference of a kind term is the kind whose tokens form the dominant source of causal origin of the body of information that S has associated with the term.¹² On Devitt's view, the reference of a kind term as used by S is the kind whose tokens ground those thoughts of S that dispose S to use the term (Devitt 1981). Both of these views put descriptive content back into the picture. But on these views, descriptive content does not determine reference. Instead, Evans and Devitt regard descriptive content as one of the relata of the reference-establishing causal relation—but it is the other relatum of this relation that is the referent. And as a causal relation is not one of fit between contents, their views allow us to retain Kripke's insight that people's beliefs about the referents of the terms they use can be massively erroneous.¹³ To mark the distinctive roles of descriptive content in descriptivist theories and in the Evans–Devitt hybrid theory, I will say that according to the former, descriptive content *determines* reference, whereas according to the latter, descriptive content *sustains* reference.

¹¹ See e.g. Evans' (1973) famous 'Madagascar' case, or Devitt's (1981) discussion of the fictional natural kind term 'grugru' (pp. 191–195). See author (ms) for an empirical survey about folk intuitions regarding reference change.

¹² I say 'Evansian' because Evans did not formulate this account for kind terms but for proper names.

¹³ See also Evans's discussion of ill-grounded beliefs (Evans 1992, Ch. 5.4.).

This feature of their view allows us to account for reference change. Here, the basic idea is the following: the beliefs one associates with a term can, and often do, have multiple causal sources; moreover, these causal sources shift when new beliefs are acquired (or old ones dropped), or when new experiences provide one's beliefs with new causal sources. For this reason, what is now the main causal source of the beliefs that a speaker associates with a term, need not be in the future. Notice that this feature of their view is of particular interest to those who aim to base conceptual engineering on an externalist metasemantics. Here, accounting for reference change is not just any old constraint, but rather the key issue, and Kripke's simplistic version of the causal theory has problems dealing with it. I take these considerations to lend some *prima facie* plausibility to Evans's and Devitt's theories.

3.2 Dual content accounts of concepts

Another important source of my dual content view is a strand in current debates about the nature of concepts. To get to the core motivations of this strand, let me sketch some background. The underlying assumption of philosophers' work on concepts has been that concepts are the "building blocks" of propositional attitudes like beliefs, desires, wishes, and the like. Concepts are posited to explain the productivity of thought, and they can only do that if they abide to the principal of compositionality (e.g. Fodor 1998). The main underlying assumption of psychologists' work on concepts, on the other hand, has been that concepts are what best explains various kinds of human behavior, such as categorization, typicality judgments, dispositions to draw inferences, etc (e.g. Machery 2009).

Given that the theoretical roles of concepts have been very different in philosophy and psychology, the accounts they have ended up endorsing are very different (Rey 1983; Machery 2009; Löhr 2020). As both camps have claimed to offer theories of the same phenomenon, namely that of concepts, this divide has spurred much confusion. However, by now, many attempts have been made to reconcile the two literatures (Margolis 1998; Prinz 2002; Weiskopf 2009). A particularly attractive way of doing this employs the following reasoning. Both the psychologically relevant roles and the philosophically relevant roles of concepts are important, but they cannot both be played by either of the candidate structures offered by philosophers and psychologists. So how can the two be reconciled? The first step is to clearly distinguish between *extension-determining structure*, which fixes the extension of a concept, and *cognitive structure*, which is used in psychological processes such as categorization, typicality judgments, drawing inferences and the like. The second step is to ask how, if at all, these structures are connected. And here the main idea is the following: extension-determining structure settles what a concept refers to (*its r-content*), whereas cognitive structure consists of beliefs about the members of the extension of the concept (*its c-content*), and can therefore explain how people apply the extension-determining structures to objects in the world. Theories that regard both r-content and c-content as kinds of conceptual content are sometimes labeled *dual content* theories of concepts (Margolis and Laurence 1999; Weiskopf 2009; Del Pinal 2015).

3.3 Two contents or dual contents?

Notice the parallels between the motivations for hybrid theories of reference on the one hand, and dual content theories of concepts on the other. Both views start from the observation that, at least for many terms and concepts, descriptive content cannot properly fix reference, but nevertheless come to acknowledge that descriptive content does play an explanatory role in how we actually apply (or in how we change to apply) our concepts and lexical items. This parallels the observation that neither associated descriptive content nor mere semantic meanings suffice to provide a complete picture of conceptual engineering.

Should these two contents both be regarded as *conceptual* content, though? In particular, are cognitive/descriptive contents not “merely associated with concepts rather than structures that are part of the nature of concepts” (Margolis and Laurence 1999, p. 42)? This is a difficult and controversial question. Defenders of dual content views have given various arguments for thinking that cognitive and referential contents are both core ingredients of concepts and/or linguistic meaning: Putnam argues that “stereotypes” belong to core meanings because they play an essential (and non-reducible) role in linguistic competence. Block (1987) argues that an otherwise plausible inferential role semantics needs to incorporate referential content in order to deliver truth conditions. Weiskopf (2009) argues against Fodor that dual contents are needed for a plausible individuation criterion of concepts, and that existing psychological models of concepts are already (implicitly) committed to dual contents anyway (see also Carey 2009). Del Pinal (2015) argues that non-extension determining structure “play[s] a key compositional role in certain constructions, especially in privative noun phrases” (p. 1).

Since this is primarily a paper in conceptual engineering, an in-depth discussion or even a full defense of the idea of dual contents is beyond its scope. When Block (1987) raises the question whether a two-factor view of meaning is to be preferred over a one-factor model, his answer is: “one scheme is better for some purposes, the other for other purposes” (fn. 31, p. 187). This is exactly the position that I will adopt within this paper. My key goal is to advocate a view with the potential of incorporating both the semantic and the practical goal. This view makes use of two different contents. As we shall see in Sect. 6, this view has further downstream benefits as well. My primary purpose is thus to give a comprehensive account of the subject matter of conceptual engineering and for this purpose, it is attractive to view both contents as conceptual contents. This allows us to say, for example, that conceptual engineering targets concepts, as the name suggests, and that it has a unified subject matter. All of this is compatible with the possibility that some of the traditional explananda in the philosophical literature on concepts are better served by treating only referential content as conceptual content.

4 The dual content view

Without further ado, then, let me state my preferred version of the dual content account:

Dual content view (DCV)

r-content: The extension of a subject S 's lexicalized concept C is the kind whose tokens constitute the main causal source of a particular subset of S 's mental states, namely those that dispose S to use lexicalizations of C .

c-content: The cognitive content of S 's lexicalized concept C consists of those of S 's mental states that dispose S to use lexicalizations of C .

Let me illustrate the inner workings of this view using the example of 'cow'. You hold many beliefs about the world.¹⁴ Only some of those beliefs dispose you to apply the word 'cow'. The belief that $2 + 2 = 4$ does not, for instance, whereas the belief that cows are relatively big mammals does. Beliefs of the latter kind constitute the c-content of your concept of cow. All the different bits and pieces of this c-content have causal origins. Some of the things you believe about cows are based on testimony and others are gained by direct acquaintance with cows. If the testimony you received does not involve mistakes or deceptions, then its content can ultimately be traced back to direct acquaintances with cows. We may therefore suppose that all or most of your beliefs regarding cows have cows as their causal source. If this is the case, then tokens of the type cow constitute the main causal source of the mental states that dispose you to use the word 'cow'. Hence, cows are the r-content of your cow-concept.

Four clarifications are in order. Firstly, a very general remark: DCV views concepts as abstract types of mental representations that have physical mental tokens. Moreover, different people may have concepts with the same r-content but with different c-content and *vice versa*. We may both possess a concept that refers to cows, but since you live on a farm, your interaction with cows has caused you to have cow-beliefs that are far more accurate than those I have gained from television. Therefore, we both have a concept with the same r-content, although with very different c-contents. Likewise, it could happen that someone tells you the exact same things about blue whales that somebody else tells me about Greenland whales, e.g. that they are mammals living in the ocean that can be over 40 feet long. If our two testifiers stand in (mediated) causal relations with blue whales and Greenland whales and the information we get from them is all we know about these whales, the two of us will have concepts with the same c-content, even though yours refers to blue whales whereas mine refers to Greenland whales.

Secondly, the view is restricted to lexicalized kind concepts. *Lexicalized* concepts are concepts for which there are natural language expressions. Most, if not all, concepts we possess are like this. *Kind* concepts are non-indexical referring concepts, i.e. concepts such as COW, DOG, FISH, WATER, DESK, KNOWLEDGE, WOMAN, or PARENT, but not concepts such as ME, YOU, HERE, NOW, AND, or OR. The concepts typically focused on by conceptual engineers are lexicalized kind concepts, so this restriction of the view is harmless for present purposes. What I suggest, then, is a generalization from natural kind externalism, as championed by e.g. Kripke and Putnam, to something that Haslanger (2006) calls *objective type externalism*. According to objective type externalism, "externalist insights should be applied to our thought and language

¹⁴ In this and subsequent examples I use beliefs to illustrate c-content. It is important to stress, however, that c-content plausibly includes other kinds of mental states than beliefs, such as pictorial contents, exemplar representations, or prototype structures.

about the social as well as the natural” (395).¹⁵ This holds despite the fact that social kinds are socially constructed and thus depend on our social relations. As Sawyer (2020b) aptly puts it, social kinds are “not, for all that, stipulated into existence, and their nature is neither transparent to us nor determined by us” (p. 14). Just like natural kinds, they are objective kinds that are open to empirical investigation. Concrete exemplars of social kinds can serve as causal sources for our beliefs about them in the same way that natural kinds can. Similar ideas apply to kinds that are arguably neither natural nor social, such colors, numbers, emotions, or persons (ibid.). Since there are concrete exemplars of each of these, these exemplars can serve as causal sources for our beliefs and thus determine r-content.

Thirdly, as the example of cow shows, the causal relation that holds between the object and the relevant set of the subject’s mental states need not be direct. As a matter of fact, this relation will often be mediated through testimony from other people (Fodor 1998; Margolis 1998). What makes an object the causal source of a belief is not the fact that it is the previous element in the causal chain, but rather that the causal chain terminates in that object.

Fourthly, DCV speaks of those of *S*’s mental states which *dispose S* to use a term. Subjects have many mental states, but not all of them play a role in their dispositions to use a specific term. According to DCV, the latter are not part of the relevant concept’s c-content. As mentioned above, your belief that $2 + 2 = 4$ does probably not dispose you to use the term ‘cow’, hence it is not part of the c-content of your cow concept. Now, there might be more mental states that dispose one to use a term than one would comfortably accept as being part of one’s relevant concept. This is the *delineation problem* of concepts. Every theory of concepts which construes them as beliefs or mental states must delineate the concept-constitutive subset of those mental states from the other ones. The present account is compatible with a plurality of different approaches to solving this problem, ranging from holism, to various psychological accounts, such as Machery’s (2017) default retrievability condition.

Before elaborating on the explanatory virtues of DCV in the following sections, let me state a rough suggestion about how the above account of concepts connects with linguistic meanings:

Linguistic meaning

The extension of a term *t* in a public language *L* is the kind whose tokens constitute the *dominant* causal source of those mental states that dispose speakers of *L* to use *t*.

According to this suggestion, the linguistic meaning of a term *t*—where *t* is the lexicalization of a concept *C*—is not necessarily identical to the concept expressed by a speaker.¹⁶ This is because public languages are constrained by social factors—how the word is generally used in one’s linguistic community, how the experts use the term,

¹⁵ This said, my account does come with limitations: on the face of it, it seems difficult to see how it could explain the reference of certain abstract concepts or of concepts for non-existing things. My account shares these limitations with other causal theories of content and reference.

¹⁶ This element of my view is similar to Sawyer’s view (2018, 2020a, 2020b). For discussion, see Sect. 6 below.

etc. This feature is captured in the notion of *dominance*: the meaning of a term is a matter of which, if any, causal source is the dominant one in a linguistic community.¹⁷

5 Dual content engineering

In this section, I aim to demonstrate that both kinds of conceptual content posited by DCV can be defective and thus the appropriate objects of engineering projects. The next section draws out the benefits of this view for theorizing about (and practicing) conceptual engineering.

According to DCV, concepts have both r-content and c-content. The recent literature on conceptual engineering has shown that both of these contents can be defective. R-content can be defective for theoretical and practical reasons. In scientific enterprises, we often want the concepts we use to track the structure of reality (Sider 2011). For example, in the natural sciences, we want our concepts to denote *natural kinds*. Insofar as a concept that we use has an extension that cuts across different natural kinds, or has things other than tokens of natural kinds in its extension, it can be found to be objectionable. Haslanger (2012) argues that similar considerations can be used to assess concepts in the social sciences as well, where we should aim to track *social kinds*. She holds that gender and race concepts are a case in point: to the extent that our current concept of e.g. a woman does not track the social kind that we should concern ourselves with, it is defective and worthy of improvement.

Defects of r-content are not restricted to scientific contexts, however. We may also find that concepts that are ubiquitous in our everyday communication are defective for practical or moral reasons. In this vein, some have argued that our concept of marriage is defective because it only has heterosexual couples in its extension (see Ludlow 2014 and Cappelen 2018 for discussion), or that our concept of a parent is defective, since it tracks biological progenitors rather than primary caregivers (Haslanger 2012). Of course, these are just example cases—DCV does not entail any particular normative judgments about these or other cases.

C-content can also be defective. Machery (2017), who takes concepts to consist solely of what, on my view, amounts to c-content, argues that concepts can be evaluated with respect to the quality of the inferences they underwrite: if they are generally reliable, the concept is valid; if they are not, it is invalid. On DCV, this kind of evaluation becomes a matter of fit between a concept's r-content and its c-content. If the c-content gives a good description of, or underwrites generally reliable inferences about the r-content, then the concept is valid; if not, it isn't. Weiskopf (2009) distinguishes between cases in which there is partial but incomplete overlap between r-content and c-content, cases in which there is complete overlap, and cases in which there is no overlap. He

¹⁷ Some philosophers have thought that conventions play an explanatory role here, most notably Lewis (1975, 2002), Schiffer (1972), and Grice (1991). Others have thought that what matters is overall usage rather than conventions, perhaps complicated by patterns of deference (Burge 1979). Today, most philosophers seem to have given up on the idea of finding a neat set of necessary and sufficient conditions for specifying communal meanings in terms of individual meanings (Williamson 1992; Cappelen 2018).

calls concepts of the first type *adequate*, those of the second type *fully adequate*, and those of the last type *inadequate* (141).¹⁸

I am sympathetic to Machery's and Weiskopf's diagnoses. However, it should be noted that 'validity' and 'adequacy' are notions that evaluate the relative fit between r-content and c-content. This means that even a fully valid or fully adequate concept does not necessarily have the c-content that it should have, all things considered. If the concept's r-content is highly defective, then even perfect alignment between c-content and r-content will only get you so far. What is needed in such cases is simultaneous engineering of r-content.

This brings me to the question of whether and how r-content and c-content can be engineered. The general procedure of c-content engineering consists of changing mental states, most notably beliefs. This is a familiar trait from a large variety of phenomena including those studied by developmental psychology, research on education and the philosophy of science. When we grow up, we refine and change our concepts until they reach the required level of sophistication expected of members of our society. This development is amplified by lower and higher education. On a communal level, improvement of c-content is furthered by the sciences. Sciences aim to provide progressively better explanations of phenomena, thereby motivating us to revise what we believe about the world. C-content engineering is a subspecies of this practice—one in which the mental states of interests are those that dispose us to use certain terms.

To engineer r-content is to bring about reference change. We have seen in Sect. 3 that although Kripke's original version of the causal theory has difficulties in explaining how this works, proposals along the lines of Evans (1973) and Devitt (1981) provide ways of modeling reference change via causal influences. DCV differs in some details from their accounts, but it captures this feature equally well. According to DCV, what a concept refers to is determined by what happens to be the main causal source of those mental states that dispose the concept-possessor to use lexicalizations of this concept. Typically, such mental states have multiple causal sources, some belong to the same kind, some belong to different kinds. As we undergo new experiences and communicate with each other, new causal sources find their way into our 'network', while others eventually fade. Willingly or unwillingly, this can happen to the extent that the kind whose tokens provide the main causal source for the mental states that dispose us to use a certain term changes over time. This yields reference change of the relevant concept. When this happens in the majority of a linguistic community, or a suitable subgroup thereof, the communal meaning of the term expressing the concept changes as well.

¹⁸ This said, Weiskopf argues that sometimes demands of efficiency in categorization may trump accuracy. All things considered, a fully adequate concept might not always be the best one.

6 Benefits of the dual content view

Now that my view is on the table, let me use this final section to point out some of its benefits. DCV allows us to make progress on a variety of issues that have troubled the recent literature.

Impact. DCV allows us to accommodate both the semantic and the practical goal at once, and is thereby a view of conceptual engineering on which it has the potential to live up to its advocates' expectations. As I argued in Sect. 2, the semantic goal is to bring about meaning change, whereas the practical goal is to change how people classify things, draw inferences, and use terminology. These goals are plausibly seen as tandem goals, since striving for one while neglecting the other runs the risk of causing linguistic confusion or obliviousness. Neither semanticism nor psychologism succeed in sketching a method of conceptual engineering that can capture both of these goals.

According to DCV, semantic meanings are grounded in the r-content of the concept which disposes people to apply the term in question. So to change semantic meanings, one has to 'engineer' r-content. But the view also accommodates the psychological structures that guide people's mental and linguistic behavior with respect to the category in question. These are grounded in the concept's c-content. So to change how people classify things, which inferences they draw or how they use the terms in question, one has to 'engineer' c-content. Doing conceptual engineering well requires making sure that these two contents are aligned with each other.

Distinctness. DCV shows conceptual engineering to be an activity that is distinct from ordinary theorizing. This contrasts with the dual representation view recently put forward by Sawyer (2018, 2020a), who argues that to make sense of conceptual engineering, two distinct kinds of representations are needed: linguistic meanings, which change in the process of conceptual engineering, and concepts, expressed by words, which serve as stable anchors to subject matters. Sawyer (2020b) further distinguishes between concepts and conceptions. Concepts are "mental representations that are constituents of thoughts"; a conception is "the set of beliefs a subject associates with a concept" (p. 7). Sawyer then argues that conceptual engineering can be understood as the activity of changing the conceptions that people associate with a given subject matter, where this change in conceptions leads to different application-conditions of the term expressing the concept. On this picture, conceptual engineering is a truth-oriented project that ultimately boils down to theory change (pp. 9–10).

Sawyer's view is ingenious and has a lot of explanatory power. Nonetheless, I think that DCV has a clear advantage. Conceptual engineering is usually understood as something that contrasts with ordinary scientific or philosophical theorizing. Whereas the latter aims to scrutinize the beliefs and theories we have about a given subject matter, conceptual engineering targets the concepts that figure in these beliefs and theories. In light of this, it is striking that according to Sawyer's view, we do not actually 'engineer' concepts, but simply revise our conceptions in the face of new evidence or arguments. To be sure, changing our conceptions may sometimes effect changes in how we apply terms, which, according to Sawyer, goes hand in hand with meaning change. But the whole enterprise looks very much like what scientists and philosophers have been doing all along, namely to propose progressively better theories about a given

subject matter.¹⁹ DCV, by contrast, takes seriously the conceptual engineer’s agenda of improving *concepts* rather than *theories*, and can thus explain the key difference between conceptual engineering and other forms of theorizing.^{20,21,22}

Control. One hotly debated topic in the recent conceptual engineering literature is whether philosophers, and speakers in general, possess a sufficient degree of semantic control for conceptual engineering to be a viable enterprise. For instance, Cappelen (2018) argues, on the basis of considerations from semantic externalism, that meaning change is mostly inscrutable and not within our control. If true, this would leave a large question mark over the very idea of adopting conceptual engineering as a method or activity to be applied in philosophy and elsewhere (Koch 2018; Riggs 2019; Deutsch 2020).

A further benefit of DCV is that it undermines the basis of such bleak pessimism. Engineering c-content requires people to change the mental states they associate with certain terms. I have earlier identified child development, higher and lower education, and scientific progress as plausible paradigms of this procedure. These practices are clearly possible, even actual, and to some extent they are within our control. This is not to say that changing c-content is easy—it often isn’t. Nonetheless, it seems that the overall success of educational and scientific practices shows that we possess a significant kind of control over c-content.

But according to DCV, even r-content engineering is neither inscrutable nor beyond our control. It can sometimes be a byproduct of significant changes of c-content. If the psychological structures we associate with a term change, new causal sources may find their way into our overall network, whereas others fade. On an individual level, this can gradually lead to reference change of a person’s concept. If this happens to a suitably large number of interacting people, then it can lead to a change in the communal meaning of a term. In this way, scientific revolutions or discoveries can contribute to meaning change. Moreover, r-content may also change without any actual changes in people’s c-content, e.g. when people unwittingly start applying a concept to a new (kind of) entity. Doing this can cause our psychological structures to be causally sustained by new (kinds of) entities. So according to DCV, conceptual change is not inscrutable, and to some extent within our (long-term) control (cf. Koch 2018, forthcoming).

Amelioration as revelation. Haslanger (2006) has famously argued for a view that has become known as ‘amelioration as revelation’. The basic idea of this view is that counterintuitive revisionary definitions of say race and gender concepts can, at

¹⁹ Sawyer (2020b) explicitly embraces this consequence of her view (p. 16).

²⁰ Perhaps Sawyer thinks that engineering concepts is impossible. If so, this paper can be read as an attempt to prove her wrong from a fellow-externalist perspective. If not, Sawyer’s view fails to explain what conceptual engineering really is about, according to many of its proponents.

²¹ Similar concerns apply to Ball (2020), who uses temporal externalism to argue that revisionary analyses do not actually change linguistic meanings, but only conceptions. I read Ball (2020) not as presenting a view about conceptual engineering, but rather as a sophisticated argument against its possibility. This argument is based on the controversial “temporal externalist” idea that meanings are partly determined by future uses and discoveries. For reasons of space, however, I cannot discuss this view in more detail here.

²² Does a similar objection also apply to psychologism? Not necessarily. Since psychologism, at least in the version defended by Machery (2017), suggests that conceptual engineering targets a particular subset of our belief-like states, namely those that are retrievable by default, it is not the case that it treats just any kind of theory change as conceptual engineering.

least in some cases, be justified on the grounds that they capture the content of the concepts that we already employ in our everyday life. Haslanger models such cases on a distinction between manifest and operative concepts (Haslanger 2000). However, depending on your view of the objects of conceptual engineering, it isn't clear that aligning one's manifest concept with the operative concept should be viewed as a form of conceptual engineering (Cappelen 2018). Insofar as her proposed strategy implies merely discovering the meanings that have been there all along, there doesn't seem to be much engineering going on (p. 80).

This criticism is sound if one takes conceptual engineering to concern only linguistic meanings. However, DCV offers a straightforward and distinctive sense in which revelation can constitute an instance of conceptual engineering. Suppose that, according to some assessment, the meaning of an expression is as it should be. Nevertheless, many people are ignorant of the term's true meaning and apply it in ways that are not consistent with the term's actual meaning. Using DCV, we can describe this case as one in which the concept that disposes people to use the term in question already has the correct r-content, but not the correct c-content. So, on DCV, we can sometimes *ameliorate* c-content by *revealing* r-content. This view thus offers the resources to accommodate a whole range of cases that single content views, such as semanticism or psychologism, struggle with.

Strawson's challenge. Many authors in the conceptual engineering debate worry that ameliorating concepts might lead to changes of topic, or subject matter, rather than offering us better ways to talk and think about an issue (Haslanger 2000; Prinzing 2017; Cappelen 2018; Sawyer 2020a; Nado 2019). The problem is, roughly, that by revising our concept of, say, knowledge, we will not end up with an improvement of our old concept of knowledge, but rather with an unrelated new concept that cannot be used to address the questions that we had regarding our original concept of knowledge. This has led many conceptual engineers to think that in the process of engineering a concept, we need to preserve some kind of continuity: something that would warrant treating the successor concept as an improvement of its predecessor.

For reasons I elaborate on elsewhere, I am not particularly moved by this worry (Koch ms). Nonetheless, DCV can potentially help to answer it. Given that this view makes use of two different kinds of content, preserving one of them while simultaneously revising the other can give us the required sort of continuity. Suppose, for instance, that we are dealing with the sort of cases I earlier identified as amelioration as revelation. In such cases, we ameliorate c-content while holding r-content fixed. Preserving r-content guarantees that we are still talking and thinking about the same things, even though our dispositions to use the word have changed. In other cases, it might be the other way around: mostly preserving c-content while changing r-content can ensure that the concept in question continues to fulfill its original function in our cognitive practices. I recognize that a proper response to Strawson's challenge requires more than the short remarks just made. Nonetheless, it is worth highlighting that the general structure of DCV can plausibly help to develop such a response.

7 Conclusion

Conceptual engineers typically endorse both semantic and practical goals, but the two leading views of conceptual engineering—semanticism and psychologism—fail to offer any comprehensive means of achieving them. The dual content view outlined in this paper draws on theory-independent considerations to construe concepts as having both referential and cognitive contents. Adopting this view allows us to construe conceptual engineering as an activity with the potential to achieve both goals at once, thereby avoiding the dangers of linguistic confusion and obliviousness. The view also provides valuable resources for making progress on a variety of hotly debated issues in the conceptual engineering literature, such as its distinctness, its feasibility, amelioration as revelation, and Strawson’s challenge.

Acknowledgements Earlier versions of this material were presented at ConceptLab in Oslo, at the EXTRA research colloquium in Bochum, at BW 11 in Barcelona, at PLM 5 in St Andrews, and at CLMPST 16 in Prague. I thank the audiences of these events for helpful feedback. Special thanks goes to Matti Eklund, Thomas Grundmann, Joachim Horvath, Manuel Gustavo Isaac, Sigurd Jorem, David Liebesman, Mark Pinder, and Sarah Sawyer for commenting on earlier versions of this paper. My work on this paper was supported by a grant of Deutsche Forschungsgemeinschaft for the Emmy Noether Independent Junior Research Group ‘Experimental Philosophy and the Method of Cases: Theoretical Foundations, Responses, and Alternatives’ (EXTRA), Project Number 391304769.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ball, D. (2020). Revisionary analysis without meaning change (or, could women be analytically oppressed?). In H. Cappelen, D. Plunkett, & A. Burgess (Eds.), *Conceptual engineering and conceptual ethics* (pp. 35–58). Oxford: Oxford University Press.
- Block, N. (1987). Functional role and truth conditions. *Proceedings of the Aristotelian Society*, 61, 157–181.
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, 4, 73–121. <https://doi.org/10.1111/j.1475-4975.1979.tb00374.x>.
- Burge, T. (1986). Individualism and psychology. *The Philosophical Review*, 95(1), 3–45.
- Burgess, A., & Plunkett, D. (2013). Conceptual Ethics I. *Philosophy Compass*, 8, 1091–1101.
- Cappelen, H. (2018). *Fixing language. An essay on the foundations of conceptual engineering*. Oxford: Oxford University Press.
- Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.
- Del Pinal, G. (2015). Dual content semantics, privative adjectives and dynamic compositionality. *Semantics & Pragmatics: A Journal of the Linguistic Society of America*, 8(7), 1–53.
- Deutsch, M. (2020). Speaker’s reference, stipulation, and a dilemma for conceptual engineers. In: *Philosophical studies*. <https://doi.org/10.1007/s11098-020-01416-z>.
- Devitt, M. (1981). *Designation*. New York: Columbia University Press.

- Eklund, M. (2015). Intuitions, conceptual engineering, and conceptual fixed points. In C. Daly (Ed.), *Palgrave handbook of philosophical methods* (pp. 363–385). Basingstoke: Palgrave Macmillan.
- Evans, G. (1973). The causal theory of names. *Aristotelian Society Supplementary*, 47, 187–225. <https://doi.org/10.1093/aristoteliansupp/47.1.187>.
- Evans, G. (1992). *The varieties of reference*. Oxford: Oxford Clarendon Press.
- Fassio, D., & McKenna, R. (2015). Revisionary epistemology. *Inquiry*, 58(7–8), 755–779.
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong (Oxford cognitive science series)*. Oxford: Clarendon.
- Glüer, K., & Wikforss, A. (2009). The normativity of meaning and content. In Zalta, E. N. (Ed.), *The stanford encyclopedia of philosophy* (fall 2020 edition). <https://plato.stanford.edu/archives/fall2020/entries/meaning-normativity/>.
- Grice, H. P. (1991). *Studies in the way of words* (1st ed.). Cambridge, Mass: Harvard University Press.
- Haslanger, S. A. (2000). Gender and race: (what) are they? (What) do we want them to be? *Nous*, 34, 31–55. <https://doi.org/10.1111/0029-4624.00201>.
- Haslanger, S. A. (2006). What good are our intuitions? *Aristotelian Society Supplementary*, 80, 89–118. <https://doi.org/10.1111/j.1467-8349.2006.00139.x>.
- Haslanger, S. A. (2012). *Resisting reality: Social construction and social critique*. New York: Oxford University Press.
- Haslanger, S. A. (2020). How not to change the subject. In T. Marques & A. Wikforss (Eds.), *Shifting concepts: The philosophy and psychology of conceptual variability*. Oxford: Oxford University Press.
- Isaac, M. G. (2020). How to conceptually engineer conceptual engineering? *Inquiry*. <https://doi.org/10.1080/0020174X.2020.1719881>.
- Koch, S. (2018). The externalist challenge to conceptual engineering. *Synthese*. <https://doi.org/10.1007/s11229-018-02007-6>.
- Koch, S. (ms). Why conceptual engineers should not worry about topics.
- Koch, S. (forthcoming). There is no dilemma for conceptual engineering. Reply to Max Deutsch. *Philosophical Studies*.
- Kripke, S. (1970/1980). *Naming and necessity*. Oxford: Basil Blackwell Ltd.
- Lewis, D. (1975). Languages and language. In K. Gunderson (Ed.), *Language, mind, and knowledge* (pp. 3–35). Minneapolis: University of Minnesota Press.
- Lewis, D. (2002). *Convention: A philosophical study*. Oxford: Blackwell.
- Löhr, G. (2019). Social constructionism, concept acquisition and the mismatch problem. *Synthese*. <https://doi.org/10.1007/s11229-019-02237-2>.
- Löhr, G. (2020). Concepts and categorization: Do philosophers and psychologists theorize about different things? *Synthese*, 197(5), 2171–2191. <https://doi.org/10.1007/s11229-018-1798-4>.
- Ludlow, P. (2014). *Living words. Meaning underdetermination and the dynamic lexicon*. Oxford: Oxford University Press.
- Machery, E. (2009). *Doing without concepts*. Oxford, New York: Oxford University Press.
- Machery, E. (2017). *Philosophy within its proper bounds*. Oxford: Oxford University Press.
- Manne, K. (2018). *Down girl. The logic of misogyny*. Oxford: Oxford University Press.
- Margolis, E. (1998). How to acquire a concept. *Mind and Language*, 13, 347–369. <https://doi.org/10.1111/1468-0017.00081>.
- Margolis, E., & Laurence, S. (1999). *Concepts: Core readings*. Cambridge, Mass: MIT Press.
- Nado, J. (2019). Conceptual engineering, truth, and efficacy. *Synthese*. <https://doi.org/10.1007/s11229-019-02096-x>.
- Pinder, M. (2019). Conceptual engineering, metasemantic externalism and speaker-meaning. *Mind*, 1, 1. <https://doi.org/10.1093/mind/fzz069>.
- Prinz, J. J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, Mass: MIT Press.
- Prinzing, M. (2017). The revisionist's rubric: conceptual engineering and the discontinuity objection. *Inquiry*, 61(8), 854–880. <https://doi.org/10.1080/0020174X.2017.1385522>.
- Putnam, H. (1975). The meaning of 'meaning'. In H. Putnam (Ed.), *Philosophical paper vol. 2. Mind, language and reality*, Philosophical papers (pp. 215–271). Cambridge: Cambridge University Press.
- Rey, G. (1983). Concepts and stereotypes. *Cognition*, 15, 237–262. [https://doi.org/10.1016/0010-0277\(83\)90044-6](https://doi.org/10.1016/0010-0277(83)90044-6).
- Riggs, J. (2019). Conceptual engineers shouldn't worry about semantic externalism. *Inquiry*. <https://doi.org/10.1080/0020174X.2019.1675534>.
- Sawyer, S. (2018). The importance of concepts. *Proceedings of the Aristotelian Society*, 118(2), 127–147.

- Sawyer, S. (2020a). Thought and talk. In H. Cappelen, D. Plunkett, & A. Burgess (Eds.), *Conceptual engineering and conceptual ethics* (pp. 379–395). Oxford: Oxford University Press.
- Sawyer, S. (2020b). Truth and objectivity in conceptual engineering. *Inquiry*. <https://doi.org/10.1080/0020174X.2020.1805708>.
- Scharp, K. (2013). *Replacing truth*. Oxford: Oxford University Press.
- Schiffer, S. (1972). *Meaning*. Oxford: Oxford University Press.
- Sider, T. (2011). *Writing the book of the world*. Oxford: Oxford University Press.
- Strawson, P. F. (1991/1963). Carnap's view on constructed systems versus natural languages in analytic philosophy: The two methods. In P. A. Schilpp (Ed.), *The philosophy of Rudolf Carnap* (pp. 502–519). La Salle/Ill: Open Court.
- Thomasson, A. L. (2020). A pragmatic method for conceptual ethics. In H. Cappelen, D. Plunkett, & A. Burgess (Eds.), *Conceptual engineering and conceptual ethics* (pp. 435–458). Oxford: Oxford University Press.
- Weiskopf, D. A. (2009). Atomism, pluralism, and conceptual content. *Philosophy and Phenomenological Research*, 79, 131–163. <https://doi.org/10.1111/j.1933-1592.2009.00269.x>.
- Williamson, T. (1992). Vagueness and ignorance. *Proceedings of the Aristotelian Society*, 66(supplement), 145–162.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.