

Change your Way of Thinking

A Neo-Fregean View on Conceptual Engineering

Cyrill Mamin

Friedrich-Schiller-Universität Jena

cyrill.mamin@uni-jena.de

Forthcoming in *Studia Philosophica*, special issue on Concepts and their Uses, ed. by Claus Beisbart.

penultimate version (April 2024)

The Neo-Fregean view individuates concepts at the level of Fregean senses. It is an internalist view according to which concepts can be described as ways of thinking that imply classifications and epistemic/normative inferences. In this paper, I argue that the Neo-Fregean view of concepts adequately characterises the targets of conceptual engineering (CE), which I depict as the activity of purposefully changing our concepts in order to change classifications and/or concept-implicit inferences. I discuss and reject rival views that either reduce the targets of CE to linguistic meanings or classification procedures or characterise concepts in a psychological or concept-externalist manner. I argue that the Neo-Fregean view is a more encompassing and adequate characterisation of the targets of CE.

1. Introduction

Conceptual revision has always been an aim in philosophy and society in general. It has been motivated by the idea that our concepts are not necessarily the best concepts we could have. Instead of reflecting on the world or trusting that our concepts are guides to truths, conceptual scepticism suggests that many of our concepts are deficient and need revision.

Conceptual engineering (CE, henceforth) is a more recent label for these activities aiming at conceptual revision. In the lively debate on CE, many views on what CE is are present,

and a thematically broad spectrum of questions are discussed.¹ One of them asks what the *targets* of conceptual engineering are, and in this paper, I will focus on this question. What do we, as conceptual engineers, aim to revise/engineer? “CE targets *concepts*, what else?” one could think. However, different alternatives have been proposed. They suggest reducing the targets of CE to something allegedly less obscure than concepts. ‘Conceptual engineering’ then could be a misnomer.² Second, even if one grants that CE targets concepts, the question remains what concepts are.

My main claims are that (1) CE is, irreducibly, targeting *concepts* and (2) the so-called Neo-Fregean view, characterising concepts as *ways of thinking*, is an adequate choice for a notion of ‘concept’ in the context of CE.

In sect. 2, I will characterise two kinds of CE that a theory on the target of CE should be able to encompass. In sect. 3, I will outline how the Neo-Fregean view on concepts fits the two kinds of CE mentioned. In sect. 4, I will discuss and reject rival views which hold that CE targets linguistic meanings or classification procedures. In sect. 5, I will critically examine two rival views on CE which include concepts but characterise concepts unlike the Neo-Fregean view (psychologically or in an externalist fashion).

2. Kinds of Conceptual Engineering

I will defend the claim that concepts/ways of thinking are the targets of CE by choosing a bottom-up approach. In contrast to starting from a general commitment to a theory of concepts, the bottom-up approach starts from the examples typically discussed under the heading of CE and then chooses a characterisation of ‘concept’ fitting these examples. I am looking for an adequate notion of ‘concept’ *for the domain of CE*, and I accept it may not be valid for other domains.

¹ See Herman Cappelen; David Plunkett: Introduction. A guided tour of conceptual engineering and conceptual ethics, in: *Conceptual Engineering and Conceptual Ethics*, ed. by Alexis Burgess, Herman Cappelen et al. (Oxford: Oxford University Press, 2020) 1–26.

² See Manuel Gustavo Isaac: What should conceptual engineering be all about?, in: *Philosophia* 49 (2021) 2053–2065, here: 2058.

Taking this route, I suggest that we have two general kinds of CE that any view about the target of CE should be able to incorporate. I will outline each kind in turn.

2.1 First kind of CE – changing classifications

Very frequently, activities of CE aim at *changing classifications*. In the socio-political domain, this kind of CE is particularly widespread, mainly concerning classifying groups of people or actions.

For *classifying people*, concepts in the domain of gender are salient; for instance, widening the concept MAN to include trans men would count as CE. Another example is the discussion revolving around the concept REFUGEE. Should people migrating for economic reasons count as refugees? People who say so demand to change the concept REFUGEE in order to include such migrants. Similarly, some activities of CE were linked to changing the institution of marriage, allowing same-sex couples to marry. Such changes in the legal institution presupposed (or, at least, went along with) changing a normative concept MARRIAGE (classifying the kinds of couples who count as ‘marriage-worthy’).

For classifying *acts*, examples with even stronger moral and political implications come to the fore. For instance, it has been debated what should count as SEXUAL HARASSMENT (e.g., only acts involving unwanted touching or verbal acts as well? If so, which kind of verbal acts?) Similarly, in earlier times, the concept RAPE generally excluded acts of rape within marriage.³ Regarding such a concept of rape (which may persist in a given society), one might wish to change how these acts are classified so that there is no difference between sexual violence outside of or within marriage.

2.2 Second kind of CE – changing implicit normative inferences

The second kind of CE aims to change concepts in order to change inferences implied by these concepts. It is particularly prominent in the case of socio-political concepts, in which I am mainly interested here.

³ Sarah Sawyer: The importance of concepts, in: Proceedings of the Aristotelian Society 118/2 (2018) 127–147, here: 143–45.

I'll illustrate the second kind of CE with an example from Sally Haslanger.⁴ Consider a conceptual engineer who proposes to replace the concept MEAT with FLESH OF TORTURED ANIMALS. This replacement will typically be expressed by proposing to redescribe a given classification. For instance, looking at the same goods in a supermarket, Harry may say "this is meat". Sally may disagree and answer "no, this is flesh of tortured animals." The engineering part would not only consist in Sally developing her new way of thinking about meat but also in trying to persuade Harry to adopt it.

Such a redescription is not mere wordplay. Each word expresses a different concept (MEAT / FLESH OF TORTURED ANIMALS), even if both concepts yield the same classification. Following Robert Brandom, the concepts' *circumstances* of application are the same, but the *consequences* of their application differ.⁵ Accordingly, concepts are *normative inference tickets*.⁶ They license drawing certain inferences (and forbid drawing others) and typically do so implicitly. For instance, thinking of something as MEAT may licence producing (involving factory farming and slaughtering), buying, cooking and eating it, whereas thinking of the same items as FLESH OF TORTURED ANIMALS may render all of these acts less legitimate, at least.

CE of this kind aims to change which inferences an individual or larger group of people *ought* to draw (rather than are actually disposed to draw) from a given conceptual classification. Depending on which concept/way of thinking is endorsed by an individual or more prominent in a society, different inferences will be more or less licenced in the view of this individual or the dominant view in society.

⁴ Sally Haslanger: Racism, ideology, and social movements, in: *Res Philosophica* 94/1 (2017) 1–22, here: 19.

⁵ Robert Brandom: *Making it explicit. Reasoning, representing and discursive commitment* (Cambridge, Mass.: Harvard University Press, 1994) 125–130. Brandom takes this difference from Michael Dummett: *Frege's philosophy of language* (New York: Harper and Row, 1973).

⁶ Jen Foster; Jonathan Ichikawa: Normative Inference Tickets, in: *Episteme* 43 (2023) 1–27.

3. Concepts/ways of thinking as the targets of CE

Returning to the question about the targets of CE, we need a notion of concepts which is adequate for both kinds of CE mentioned. My aim is to show that the Neo-Fregean view on concepts is particularly apt for fulfilling this task.

3.1 *The Neo-Fregean view – concepts as ways of thinking*

To begin with, I'll briefly explain what Fregean senses are. Frege introduces the notion of 'sense' related to linguistic expressions (or 'signs', more generally). He distinguishes between "that to which the sign refers, which may be called the reference of the sign" and "the *sense* of the sign, wherein the mode of presentation is contained".⁷ Each sign has a determinate sense and, thereby, a determinate reference. At the same time, different signs (expressing different senses) can have the same referent.⁸ For instance, one single creature (or a class/kind of creature) may either only be cognitively presented as a creature with a kidney ('renate') or as a creature with a heart ('cordate').⁹

The Neo-Fregean view *individuates concepts at the level of Fregean senses*.¹⁰ Christopher Peacocke's criterion of distinctness (of two concepts) is a version of such an individuation.

⁷ Gottlob Frege: On Sense and Reference, in: *Translations from the Philosophical Writings of Gottlob Frege*, ed. by Peter Geach and Max Black (Oxford: Basil Blackwell 1952) 56–78, here: 57.

⁸ *Ibid.*, 57–58.

⁹ The notion of 'presentation' is still quite vague, given that there are presentations (e.g. visual ones) which have nothing to do with concepts. Peacocke's and Evans's explications of Fregean sense will make things more precise (see below).

¹⁰ The Neo-Fregean view is 'neo' mainly because Frege himself did not identify concepts with senses. According to Frege, concepts are functions, and as such, they are the *referents of predicates*. See Gottlob Frege: *Comments on Sense and Meaning*, in: *Posthumous Writings*, transl. by Peter Long and Roger White (Oxford: Basil Blackwell, 1979) 118–125, here: 118; Gottlob Frege: *On Concept and Object*, in: *Translations from the Philosophical Writings of Gottlob Frege*, ed. by Peter Geach and Max Black (Oxford: Basil Blackwell, 1952) 42–55, here: 43(fn.).

Concepts *C* and *D* are distinct if and only if there are two complete propositional contents that differ at most in that one contains *C* substituted in one or more places for *D*, and one of which is potentially informative while the other is not.¹¹

To illustrate, compare the sentences (A) ‘a renate is a renate’ and (B) ‘a renate is a cordate’. The propositional contents expressed by these sentences differ at most in that in (B), the concept RENATE is substituted with the concept CORDATE in one place, which renders only (B) potentially informative.

By referring to the criterion of *informativeness*, concepts are individuated at the level of ‘cognitive significance’, where cognitive significance must be specified from the perspective of subjects possessing the concepts.¹² After all, content is informative (or not) for *someone*. For instance, Neo-Fregeans hold that RENATE and CORDATE are two different concepts, even though they have the same extension. Individuating concepts in this way comes close to widespread pre-theoretical meanings of ‘concept’.¹³ Given the bottom-up approach I’ve chosen, this connectivity to everyday concept talk is an important advantage of the Neo-Fregean view.

The fact that the Neo-Fregean view individuates concepts in line with the cognitive perspectives of concept-users makes it a so-called *internalist* view of concepts, which is opposed to both concept-externalist and subjectivist/psychological views. What are the differences? *Concept-externalists* claim that concepts are individuated externally, e.g. by causal links to the concepts’ referents (I will return to concept-externalism in sect. 5.2). *Subjectivists* characterise concepts as ‘mental representations’, which undermines the Fregean difference between senses and subjective imaginations (I will return to subjectivism in sect. 5.1).

This seems to be a more restricted notion of ‘concept’, according to some Neo-Fregeans, see e.g.: Christopher Peacocke: *A Study of Concepts* (Cambridge, Mass.: MIT Press, 1992) 2. Nonetheless, it can be argued that this is rather a matter of terminology and Frege would, ultimately, have agreed to identify concepts with senses (thanks to an anonymous reviewer for this remark).

¹¹ C. Peacocke: *A Study of Concepts*, op. cit., 2.

¹² *Ibid.*, 3.

¹³ See also Hans-Johann Glock: *Concepts. Where subjectivism goes wrong*, in: *Philosophy* 84/1 (2009) 5–29, here: 10.

Individuating concepts at the level of Fregean senses becomes more plausible by explicating senses as ways of thinking, as Gareth Evans does.

I suggest that we take Frege's ascription of a sense to a Proper Name to mean that not only must one think of an object—the referent of the term—in order to understand a sentence containing it, but also anyone who is to understand the sentence must think of the referent *in the same particular way*. It is therefore, for Frege, as much a public and objective property of a term that it imposes this requirement, as that it has such and such an object as its referent.¹⁴

In this passage, Evans writes about the sense of proper names instead of predicates (or ‘*concept-words*’). Therefore, one might think that it does not fit concepts. However, reflecting on the senses of proper names is just the right thing to do. Iff a proper name *refers* to an object, the *sense* of the proper name contains the way of thinking about (in Frege’s original terms: ‘mode of presentation of’) this object. Accordingly, if concepts are individuated at the level of Fregean senses, it follows that concepts can be identified with ways of thinking about the proper name’s (or any other expression’s) referent.

Each mode of presentation/way of thinking of an expression’s (typically, a proper name’s) referent can be made explicit by expressing it in predications. To illustrate this with one of Frege’s examples: Two explorers are approaching the same mountain from different directions; one calls the mountain ‘Aphla’, and the other calls it ‘Ateb’. What makes the sentence ‘Aphla is Ateb’ informative (in contrast to ‘Aphla is Aphla’)? Frege answers that it is the *difference in sense* between ‘Aphla’ and ‘Ateb’.¹⁵ According to Evans, this can be explicated by saying that ‘Aphla’ and ‘Ateb’ correspond to two different *ways of thinking* about the same mountain.¹⁶ This implies a difference regarding the *previous predications* about the mountain. For instance, explorer A said «Aphla is the holy mountain of the Aphlans», and explorer B said «Ateb is the highest mountain of the Atebs». In

¹⁴ Gareth Evans: *The Varieties of Reference* (Oxford: Clarendon Press, 1982) 17 (italics in the original).

¹⁵ Gottlob Frege: Letter to Jourdain, in: *Meaning and Reference*, ed. by A. W. Moore (Oxford: Oxford University Press, 1993) 43–45.

¹⁶ G. Evans: *The Varieties of Reference*, op. cit., 16.

revealing that Aphla is Ateb, the explorers can unify their previous predications, which makes the identity statement truly informative.¹⁷

In what follows, I will argue that CE targets concepts, where concepts are characterised as ways of thinking in the Frege-Evans-spirit. Regarding CE, this implies that changing concepts/ways of thinking amounts to changing (implicit or explicit) predications. More precisely, “to change” here means “to replace” because one way of thinking and its associated predications is replaced with another. As I will argue, replacing such predications can have two different *aims*: to classify objects differently or to change the inferences to other concepts/ways of thinking (or both).

3.2 Ways of thinking as the targets of CE

I will try to make plausible that the Neo-Fregean view of concepts is adequate for both kinds of CE (as mentioned in sect. 2).

3.2.1 Ways of thinking in the first kind of CE (classifications)

Classifications are based on ways of thinking, that is, on how parts of the world are presented to concept users. For instance, if one thinks of people’s gender based on a ‘traditional’ way of thinking, they will classify people based on bodily appearances as MAN and WOMAN. If they think of people’s gender in a different way, the result will be a different classification, e.g. taking more gender kinds into account. CE might replace one way of thinking, implying the traditional gender classification, with a different way of thinking, implying a new gender classification.

How I explained this example might seem puzzling in the following way. The Fregean picture implies that the *same* object can be given in different ways of thinking. However, if these are different ways of thinking/concepts about just one object, one might expect

¹⁷ In contrast, if someone just learns that expression B refers to the same object as expression A, where A and B already implied the same predications about the common referent, this would not be truly informative.

these concepts to be co-extensional.¹⁸ In contrast, in the gender example, the concept-extensions typically differ (e.g. MAN excludes trans men, MAN* includes them).

In examples like gender, the seeming puzzle can be dissolved if we are more liberal concerning what x stands for in ‘way of thinking of x ’ or ‘mode of presentation of x ’. It does not have to be an entity at the level of the objects classified. Instead, there is a common *topic* (gender), and the two concepts at stake are two ways of thinking about this topic. The Neo-Fregean picture applies because it makes sense to say that the two ways of thinking are two different ways of thinking *about something*. Resultingly, concepts with different extensions can be ways of thinking about the same topic.

In other cases, concepts/ways of thinking with different extensions can even be ways of thinking of the same *object*. For example, a pre-amelioration concept RAPE excluding marital rape and a post-amelioration concept RAPE* including marital rape may both refer to the same social or moral kind *rape* (I’ll return to this case in sect. 5.2). I consider it an important advantage of the Neo-Fregean view that it *allows* (instead of requires) the pre- and post-engineering concept to be ways of thinking of the same object.

3.2.2 Ways of thinking in the second kind of CE (inferences)

The Neo-Fregean notion of concepts is also adequate regarding the second kind of CE. To see this, let’s come back to the meat example. Harry’s and Sally’s respective concepts (MEAT / FLESH OF TORTURED ANIMALS) have the same extensions. (For instance, think about the situation in the supermarket, where their respective concepts pick out the same items.) Nevertheless, it makes sense that Harry and Sally apply different concepts. The reason is that even though Harry’s and Sally’s concepts/ways of thinking have the same extension, they do not license the same implications. As a result, Harry’s and Sally’s predications may differ. For instance, if Harry predicates about something that it is meat, he feels licenced to also predicate about it that we can buy and eat it. If Sally predicates

¹⁸ Frege holds that if two concepts are co-extensional, they are, in fact, just the same concept (see G. Frege: On Concept and Object, op. cit., 118). The Neo-Fregean view, individuating concepts at the level of Fregean senses, cannot agree with this view (see also sect. 3.2.2).

about the same thing that it is flesh of tortured animals, this may lead her to resist Harry's predication.

Individuating concepts at the level of Fregean senses is just the right level for describing the targets of CE. On the one hand, the individuation is fine-grained enough, so concepts are treated as different even if they imply the same classification. This is important, for CE often aims to change predications not to change classifications but to change inferences. On the other hand, the individuation is not too fine-grained, so personal ideas associated with the same concept may still vary.

4. Rival views A) – against concepts as the targets of CE

In the remainder of this paper, I will discuss some rival views concerning the target of CE. In the present section, I will start with two views that do not take concepts to be the targets of CE. In sect. 5, I will go on with two rival views which grant concepts some place in CE but do not characterise concepts in the Neo-Fregean way I've proposed.

4.1 Linguistic meanings

The view that CE targets linguistic meanings is widespread. I will discuss two versions of the linguistic view: the view that CE targets semantic (or 'lexical') meanings and the view that CE targets *speaker meanings*.

Herman Cappelen defends the first version of the linguistic view. He claims that CE aims to change the intensions and extensions of linguistic expressions.¹⁹ More precisely, Cappelen holds that "conceptual" amelioration «*always involves the extension and intension of a predicate changing over time*»,²⁰ where the change in extension happens due to the change in intension.

Cappelen combines this view with *semantic externalism*, a view that takes the external environment of speakers to determine the extensions and intensions of expressions, at least in part. This view includes a possibility of widespread and «massive, fundamental

¹⁹ Herman Cappelen: *Fixing Language. An Essay on Conceptual Engineering* (Oxford: Oxford University Press, 2018) 61–62.

²⁰ *Ibid.*, 62 (italics in the original).

mistakes and confusions about semantic values». ²¹ In addition, there is no ‘recipe’ or ‘algorithm’ telling us how we could intentionally change semantic meanings. ²²

This leads Cappelen to a sceptical conclusion regarding the possibility of CE. The problem is particularly salient for semantic externalism, but Cappelen sees semantic internalists as equally affected. ²³ The general worry is that we do not seem to have enough control over the facts and mechanisms determining linguistic meaning. ²⁴

There have been different objections to Cappelen’s sceptical outlook on the possibility of CE. One of them, going in the direction I am interested in here, comes from Mark Pinder, who develops an alternative picture of what kind of meaning is targeted in CE. According to Pinder, it is not semantic but *speaker meaning*. This is how Pinder characterises the difference (based on H.P. Grice):

speaker meaning is what a speaker intends to mean by her words; *semantic meaning* is the meaning assigned to the word by the speaker’s language. Here, we can assume that the speaker’s language is a *public* language, perhaps such as English or French or dialects thereof. Importantly, though, one can speaker-mean *F* by *w* regardless of whether *w* semantically-means *F* in one’s language. ²⁵

To illustrate, a simple example in which speaker meaning becomes salient is the game of two children who intend to mean no by ‘yes’ and yes by ‘no’. ²⁶

The speaker-meaning picture remedies Cappelen’s metasemantical worries about the possibility of CE. Obviously, CE *is* possible from the speaker-meaning view, even though Cappelen may be right in saying that one cannot easily change a term’s lexical meanings. This is why the speaker-meaning picture approaches the view that CE targets ways of

²¹ Ibid., 63.

²² Ibid., 67.

²³ Ibid., 81–82.

²⁴ Ibid., 58–69.

²⁵ Mark Pinder: The Austerity Framework and semantic normativity, in: *Inquiry* 64/1-2 (2021) 123–141, here: 131.

²⁶ See Mark Pinder: Conceptual engineering, metasemantic externalism and speaker meaning, in: *Mind* 130/517 (2021) 141–163, also for further examples.

thinking. Plausibly, attempts to change ways of thinking often accompany speaker-meaning activities. For instance, the first activists who advocated for including trans men in the definition of ‘men’ probably were in the business of changing a speaker-meaning for a local group of speakers first. So, from the perspective of the ways of thinking approach, the speaker-meaning picture is an improved version of the linguistic view, and the two views can be combined.

However, at least two objections against the linguistic view remain, even in the speaker-meaning version. I will explain them in turn.

My *first objection* is that semantic and speaker meanings leave the target of CE underdetermined. Changed semantic meanings or speaker meanings are neither necessary nor sufficient for CE.

To explain the objection, I begin with Pinder’s example of the children’s yes-no-game. This is an instance of changed speaker meanings, but it is not an instance of CE, for nothing substantial has changed here. If one swaps the meanings of ‘yes’ and ‘no’, this decision neither promotes nor reflects a changed attitude or worldview. However, I take it for granted that CE promotes and/or reflects such changes. So, it seems that changed speaker meanings are not a sufficient condition for CE. Now, Pinder would not object here. He just says that a change of speaker meaning does not need to be sufficient for CE. This change only is what he calls the ‘descriptive’ aspect, whereas he adds that in CE, the added normative aspect is that the participants think «it is a good thing to speaker-mean x». ²⁷

My reply is that Pinder should take one step further. When is it a good thing to speaker-mean x? What is the reason why people aim to change speaker meanings? Following Pinder, there is no further theoretical story to tell here. In contrast, reflecting on the examples of CE, I think we can and should say more.

A similar objection of underdetermination, in the sense of not giving sufficient conditions for the occurrence of CE, also affects Cappelen’s version of the linguistic view. To see this, I propose to ask and answer two questions. First, can you imagine a situation in which linguistic meanings in a given domain significantly change (the lexical meanings,

²⁷ Ibid., 155.

not only speaker meanings) without affecting a person's concepts? Second, conversely, can you imagine a situation in which a person's concepts change without any change of linguistic meanings taking place? If so, changing linguistic meanings might not even be necessary for CE.

I think that the answer to both questions is 'yes' and that this is reason enough to say that concepts cannot be reduced to linguistic meanings as the targets of CE.

To illustrate my claim, I can make use of a real example. Recently, a bishop in Switzerland – who was sceptical regarding a new law allowing same-sex couples to marry – said that one should invent a new term, 'bio-marriage', to refer to the kind of marriage that excludes same-sex couples.²⁸ Probably, in the society the bishop lives in, the inclusive meaning of 'marriage' has become the lexical meaning. In order to successfully communicate, the bishop just needs to accept the current lexical meaning. However, considering the first question above, the bishop did not abandon his previous way of thinking about marriage. Instead, he just uses the toolkit of the present-day language, with its actual lexical meanings, to express his old concept/way of thinking. Thereby, the bishop now *masters* two concepts/ways of thinking of marriage but he only *endorses* his previous marriage concept. Given that CE is about changing ways of thinking, there must be a change in the concepts a subject *endorses* (instead of just *masters*) for CE to be successful. What the bishop did leads us to the second question above. He just used the toolkit of the present-day language in a creative way. Imagine that the bishop convinced someone who was in favour of same-sex marriage of his own view that different-sex marriage (to which he refers with the expression 'bio-marriage') is the 'genuine' kind of marriage. If so, the bishop makes a person *endorse* his favoured way of thinking and thus to change her concept of marriage.²⁹ However, this does not amount to a change in linguistic meanings.

²⁸ Joseph Bonnemain, quoted in Erich Aschwanden; Simon Hehli: «Die Kirche sieht die Sexualität als etwas Herrliches», in: Neue Zürcher Zeitung (6.7.2021) 8–9.

²⁹ On the view proposed, changing a way of thinking can be successful on the micro-level of, say, a relationship as well as on the macro-level of an entire society.

The bishop example shows that present linguistic meanings do not completely constrain language users in expressing and propagating their ways of thinking. They may just use the expressions in different combinations or create new expressions from present ones, like ‘bio-marriage’. As a result, changing concepts does not depend on changing linguistic meanings.

My *second objection* against the linguistic view is that it does not encompass the second kind of CE (changing concepts in order to change inferences). To see this, consider again the meat example. In this example, no linguistic meaning has changed, and there was no attempt to do so. A person like Sally, who tries to convince Harry that «this is not meat but flesh of tortured animals», is not in the business of changing linguistic meanings. The terms she uses to express the new way of thinking have the meanings they already had before the process of CE started.

In sum, concepts cannot be reduced to linguistic meanings as the targets of CE. A change of linguistic meanings is not what happens in many cases of CE.

4.2 Classification procedures

More briefly, I will say something about a recent suggestion concerning the target of CE Jennifer Nado has proposed.³⁰ According to Nado, the targets of CE are *classification procedures*. Classification procedures are recipes for picking out entities. They work like algorithms (another analogy Nado uses). So, for instance, the aim for CE might be to replace a classification procedure excluding trans men with a classification procedure including them.

As Nado mentions herself, classification procedures come close to Fregean senses, so there is no huge gap to the view that concepts, characterised as ways of thinking, are the targets of CE. (For instance, the procedure of picking out renates and the procedure of picking out cordates are different, even if they result in the same extension.) Accordingly, Nado does not deny that CE might target concepts. However, Nado thinks that talking

³⁰ Jennifer Nado: Classification procedures as the targets of conceptual engineering, in: *Philosophy and Phenomenological Research* 106 (2023) 136–156.

about ‘senses’ «does nothing but muddy the waters here, tying us to various presuppositions and assumptions that are more likely to confuse than illuminate». ³¹ So, according to Nado, we do not need to refer to the contested notion of concepts. Instead, we can think about classification procedures.

I think that Nado is right that CE is often about changing classifications instead of changing linguistic meanings. However, my main worry against her view is that *we cannot explain CE of the second kind* by restricting ourselves to classification procedures. As we have seen in the meat example, the classification stays the same even though MEAT is replaced with FLESH OF TORTURED ANIMALS. So, changing a way of thinking can be an aim of CE, even though this does not change a classification.

Relatedly, reducing the targets of CE to classification procedures would not do justice to the fact that CE is an activity which involves rational subjects. Nado holds that CE does not involve persuasion; it is just designing a new classification procedure, in analogy to writing a new algorithm. ³² In contrast, saying that ways of thinking are the target of CE implies that rational subjects, having their ways of thinking, are addressed.

A remaining worry might be that ‘concept-talk’ is obscure. Recall Nado saying that talk of Fregean senses ‘muddies the waters’. My answer to this worry is simply that, as it seems, we *need* the notion of concepts and that it can be made clear enough, as I have tried to do.

5. Rival views B) – different notions of concepts in CE

In this final section, I will discuss two views on CE which include concepts, like the Neo-Fregean view does. However, these views have different notions of what concepts are. My claim will be that the Neo-Fregean account has significant advantages compared to these views when it comes to describing the targets of CE.

³¹ Ibid., 143.

³² Ibid., 147–148.

5.1 Psychological notions of concepts in CE

For the psychological or subjectivist notions of concepts, I choose Manuel Gustavo Isaac as a representative.³³ Isaac's contrast between philosophical and psychological theoretical approaches to the notion of 'concept' is instructive. According to Isaac, philosophical approaches typically take concepts to «consist of sets of separately necessary and jointly sufficient features, which are further characterised as semantically analytical and epistemically a priori».³⁴ In contrast, psychological approaches typically see concepts as «structured bodies of information about some category of referents», which are retrieved and activated in cognitive processes. On this view, concepts can be realised by different cognitive structures, e.g. exemplars, stereotypes or theories.³⁵

Regarding CE, Isaac holds that psychological theories of concepts are superior to their philosophical rivals. He claims that psychological theories do very well in both saying what is problematic about concepts and how one could change «people's minds and ways of thinking when they are found to be improvable», thereby bringing about «the desired changes in people's cognitive behavior and abilities».³⁶ In contrast, according to Isaac, philosophical theories of concepts are somehow remote from the actual conceptual practices and would be restricted to the 'most sophisticated' concepts.³⁷

Isaac's claim that theories on CE should concern «people's minds and ways of thinking» is similar to the Neo-Fregean view. However, the Neo-Fregean view is not a psychological view of concepts in Isaac's sense. At first sight, such a psychological view may seem attractive for CE due to the seeming contrast with abstract and too sophisticated philosophical accounts of concepts. However, the way Isaac pictures this contrast is flawed, in my view, for two reasons.

³³ Manuel Gustavo Isaac: Which concept of concept for conceptual engineering?, in: *Erkenntnis* 88 (2023) 2145–2169.

³⁴ *Ibid.*, 2152.

³⁵ *Ibid.*

³⁶ *Ibid.*, 2155.

³⁷ *Ibid.*, 2154.

First, Isaac has an overly narrow view of ‘philosophical’ accounts of concepts when he says that concepts must be described by necessary and jointly sufficient conditions on these accounts. This claim is called ‘descriptivism’. The Fregean view of concepts is not bound to it. For instance, there may be vague concepts and/or concepts whose conditions of application remain opaque to concept users.³⁸

Second, and more importantly, Isaac remains silent about some disadvantages of the psychological accounts, which are illuminated by considering the Fregean contrast between senses and private ideas. This is how Gareth Evans describes the distinction:

To take an example of Frege’s we must say that someone who thinks of a horse as the horse ridden by the Queen is thinking of the horse in the same way as someone who thinks of it as the Queen’s steed, for the difference in poetic colouring could never be the basis, for someone who fully grasped both senses, for taking different attitudes towards the two thoughts.³⁹

So, in contrast to more subjectivist/psychological views on concepts, CE targeting concepts as ways of thinking targets something *more coarse-grained than psychological states* (e.g. stereotypes).⁴⁰ This makes room for a more liberal kind of CE because suc-

³⁸ Frege at some point says that concepts must not have «vague boundaries» but that it «must be determinate for every object whether it falls under a concept or not» (G. Frege, Comments on Sense and Meaning, op. cit., 122). However, from the context of this remark, we can interpret this as a *specific* requirement, only valid for *logic*, in which Frege is mainly interested.

³⁹ G. Evans: The Varieties of Reference, op. cit., 120.

⁴⁰ However, we have to lower the bar for what already counts as a way of thinking/concept in contrast to ‘poetic colouring’. The question is whether two different expressions refer to the same concept (= mere difference in colouring) or refer to different concepts (= difference in way of thinking/sense). In the example Evans mentions (‘horse’/‘steed’), I’m not sure. However, take another example Frege gives for a mere difference in colouring, the difference between ‘dog’ and ‘cur’ (‘cur’ is a pejorative term for a dog; see Gottlob Frege: Logic, in: Posthumous Writings, transl. by Peter Long and Roger White [Oxford: Basil Blackwell, 1979] 140). *Pace* Frege, I think that the two expressions differ in sense and thereby refer to different ways of thinking/concepts about members of the species *canis familiaris*. My evidence is that DOG and CUR license different

cessfully changing a way of thinking does not try to change psychological states and dispositions like associations, stereotypes and things alike. Such changes may accompany changes in ways of thinking, but they are not the primary target of CE.

An example regarding *stereotypes* may highlight the difference. What comes to mind when you hear or read the expression ‘engineer’? Many people may spontaneously think of a male person, according to the still widespread gender-biased stereotype. Is *this* characterising the concept ENGINEER they possess? A proponent of the psychological view is inclined to affirm this.

In contrast, as a proponent of the Neo-Fregean view, I’m not. CE is not about changing what first comes to your mind when you hear ‘engineer’. Instead, if a person fully affirms that engineering is not exclusively for men, we can ascribe the corresponding gender-open concept to them, even though they may still be affected by the male-biased stereotype.

5.2 *The externalist notion of concepts in CE*

The Neo-Fregean view is an *internalist* view of concepts; it says that concepts are individuated by taking the perspective of the subject possessing the concept. In contrast, many philosophers are externalists about meaning and concepts.

To outline concept-externalism, I quote Sarah Sawyer.

According to the externalist, concepts are not individuated by individual conceptions—they are not individuated by the way the individual thinker takes the world to be. Nor, according to the externalist, are concepts individuated by communal conceptions [...]. Concepts are individuated by relations to objective properties independently of our individual or communal conceptions of them.⁴¹

predications about these members (e.g. being curs, it may be ok to kick them, being dogs, it may not).

⁴¹ S. Sawyer: *The Importance of Concepts*, op. cit., 139.

According to the concept-externalist, conceptions of a concept may be incomplete or even wrong. A patient who has wrong beliefs about arthritis can still be said to share the concept ARTHRITIS with, e.g. their doctor.⁴²

Classical examples in favour of externalism concern *natural kind* concepts. If we consider *social kinds* instead, concept-externalism becomes more controversial. For instance, Sawyer says that CE «brings about a change to the extension of the relevant term while maintaining continuity of reference to the same social kind over time».⁴³

I agree with Sawyer that there are cases in which we need sameness of reference throughout the engineering process. Plausibly, the pre-engineering concept RAPE (excluding marital rape) and the post-engineering concept RAPE* (including marital rape) are two ways of thinking about the same object, which is *rape*. This allows saying that the new concept more adequately represents what rape *really is*. However, Sawyer frames the example slightly differently. In her view, the concept RAPE stays the same pre- and post-engineering. What happens is that people at t1 and t2 have different *conceptions* of the concept rape.⁴⁴ Accordingly, Sawyer says that CE is not targeting concepts but conceptions.

In the ‘realist’ cases, in which we have a common referent, almost everything I said about concepts should be translatable to ‘conceptions’ in Sawyer’s framework.⁴⁵ However, Sawyer’s view may not generalise to other instances of CE. Allegedly, changing a classification often *replaces* social kinds. For instance, thinking of people’s gender in a new

⁴² See Tyler Burge: Individualism and the Mental, in: Midwest Studies in Philosophy 4/1 (1979) 73–121.

⁴³ Sarah Sawyer: Kinds of Kinds. Normativity, Scope and Implementation in Conceptual Engineering, in: New Perspectives on Conceptual Engineering, ed. by Manuel Gustavo Isaac, Steffen Koch et al. (Synthese library, forthcoming).

⁴⁴ S. Sawyer: The Importance of Concepts, op. cit.; see also Sarah Sawyer: Talk and Thought, in: Conceptual Engineering and Conceptual Ethics, ed. by Alexis Burgess, Herman Cappelen et al. (Oxford: Oxford University Press, 2020) 379–395.

⁴⁵ An exception concerns the relation between concepts/conceptions and linguistic meaning. Sawyer’s targets of CE, conceptions, are firmly connected to linguistic meaning (see S. Sawyer: The Importance of Concepts; Talk and Thought, opt. cit.), whereas my targets of CE, concepts as ways of thinking, are not.

way will typically result in a *new* social kind. One way of thinking has been replaced with another; that is, one concept WOMAN_{TRAD} has been replaced by WOMAN_{INCL}. Plausibly, each way of thinking/concept, through its extension, has *generated* a social kind (the social kinds *woman_{trad}* / *woman_{incl}*).

Now, *if a social kind has already been generated*, our aim may be to represent this pre-existing social kind, indeed. It seems to me that these are the cases Sawyer has in mind. For instance, she holds that

it is an empirical discovery that race is not, as once thought, a biological category, and that it affects, in either a positive or a negative way, one's capacity to enter the professional workforce, the probability that one will suffer lifelong systematic discrimination, and one's physical and mental health. The same can be said of warfare, poverty, crime, punishment, gender, and so on: their natures are not stipulated but discovered.⁴⁶

However, there has been a time in which (at least many of) the social kinds Sawyer mentions have been *generated* by our concepts. For instance, without human race and gender concepts, there would be no social kinds like *race* and different gender kinds. CE is often interested in *this* connection between concepts and the kinds generated by them. In these cases, we do not have a stable referent throughout the engineering process. Rather, CE aims then to re-generate, delete, or newly generate social kinds. Therefore, changing a way of thinking *may* aim at grasping the same object or kind more adequately, but changing a way of thinking may also aim at changing an object or kind.

6. Conclusion

I have argued that the targets of conceptual engineering (CE) are concepts, characterised as ways of thinking which are individuated at the level of Fregean senses. This Neo-Fregean view of concepts is adequate regarding the two kinds of CE we should be interested in, namely, (1) CE changing classifications and (2) CE changing inferences. In contrast, views that propose to reduce the targets of CE to either linguistic meanings or classification procedures fail because they are inadequate or not comprehensive regarding the

⁴⁶ Sarah Sawyer: Truth and objectivity in conceptual engineering, in: *Inquiry* 63/9-10 (2020) 1001 – 1022, here: 1014.

two kinds of CE mentioned. Also, two views that take concepts seriously but characterise them in a psychological/subjectivist or externalist way have serious disadvantages in theorising about CE if compared to the Neo-Fregean view. Claiming that CE targets ways of thinking implies that CE is a discursive process, addressing public and rational commitments of concept users who still remain flexible in their use of language and diverse in their psychological mindsets.⁴⁷

47 With thanks to Sarah Sawyer, Claus Beisbart, two anonymous reviewers for this journal, and the audiences at the conferences *Concepts and their Uses* (Bern), *PhilLang* (Lodz) and the *Collegium Philosophicum* at the University of Jena (all 2023) for helpful comments on earlier versions of this paper.