Abstract. Conceptual engineering involves revising our concepts. It can be pursued as a specific philosophical methodology, but is also common in ordinary, non-philosophical, contexts. How does our capacity for conceptual engineering fit into human cognitive life more broadly? I hold that conceptual engineering is best understood alongside practices of conceptual exploration, examples of which include conceptual supposition (i.e., suppositional reasoning about alternative concepts), and conceptual comparison (i.e., comparisons between possible concept choices). Whereas in conceptual engineering we aim to change the concepts we use, in conceptual exploration, we reason about conceptual possibilities. I approach conceptual exploration via the linguistic tools we use to communicate about concepts, using metalinguistic negotiation, convention-shifting conditionals, and metalinguistic comparatives as my key examples. I present a linguistic framework incorporating conventions that can account for this communication in a unified way. Furthermore, I argue that conceptual exploration helps undermine skepticism about conceptual engineering itself.

Keywords: Conceptual engineering, metalinguistic negotiation, counterconventional conditionals, metalinguistic comparatives

1 Introduction

Conceptual engineering is the practice of modifying and revising the concepts we use. It can be pursued as a specific methodology for resolving or diffusing philosophical debate. For instance, Plunkett (2015) argues that large parts of the free will debate are best viewed as debates over what concepts to use. Thomasson (2007, 2013, 2017) argues similarly for a range of disputes in metaphysics. Philosophers also undertake projects of conceptual engineering in the service of moral, social and political goals. Here we find, for example, Haslanger’s (2000, 2005) “ameliorative projects” surrounding race and gender terms, as well as Manne’s (2017) project arguing for a concept of misogyny that focuses on power structures rather than individual psychology. Conceptual engineering often also takes place outside of academic philosophical discourse, though it isn’t always explicitly recognized as such. Cantalamessa (2021) argues that disability rights activists are engaged in conceptual engineering when promoting a social rather than medical approach to
disability. Black activist Kennedy Mitchum contributed to conceptual engineering when she influenced *Merriam-Webster* to edit their entry for *racism* to highlight systemic oppression (Zimmer, 2020). Also engaged in conceptual engineering were the astronomers who decided, in 2006, to redefine *planet* so that Pluto no longer counts. Die-hard Pluto fans, for their part, continue to promote a countervailing concept of planet inclusive of Pluto.

In all of these cases of conceptual engineering, or attempted conceptual engineering, the idea is that we don’t *discover* what is free will, or what is racism, or what is a planet. Instead, we negotiate and choose our concepts, and thereby in an important sense *decide* these matters (which is *not* to say that all decisions are equally good).

There is, of course, substantial debate about many aspects of this process. However, my starting question is not about the details of conceptual engineering itself, but rather about how our capacity for conceptual engineering fits into human cognitive life more broadly. I hold that we gain a better understanding of conceptual engineering by situating it alongside a general practice of what I’ll call *conceptual exploration*. We explore our concepts by considering what would follow if we changed them in some way — what I’ll call *conceptual supposition*. We also explore our concepts by comparing alternative conceptual choices — what I’ll call *conceptual comparison*. While I’ll focus here on these two modes of conceptual exploration, there are surely many others that will remain to be examined. Whereas conceptual engineering changes or aims to change our concepts, conceptual exploration reasons about conceptual possibilities. Conceptual exploration is also distinct from *conceptual ethics*, or normative inquiry about what our concepts should be like (Burgess and Plunkett, 2013a,b). Still, conceptual exploration may help us settle on views in conceptual ethics, which we may then take as the goals of conceptual engineering projects. We gain a fuller understanding of our conceptual activities by broadening our view to include conceptual exploration, in addition to the more familiar subjects of conceptual engineering and conceptual ethics.

In this paper, I approach conceptual exploration via the linguistic mechanisms we use to communicate about concepts. By looking at how we communicate, not only about what the world is like factually, but also about what concepts we use to categorize things in the world, we will find that a full account of conceptual exploration places meaningful constraints on semantic the-
ory. Because of this, even skeptics about conceptual engineering cannot ignore alternative concepts when theorizing about communication. But furthermore, an examination of conceptual exploration provides abductive support for the possibility of conceptual engineering itself.

I begin my discussion of communication about concepts in section 2 by looking at disputes about concepts: cases of so-called metalinguistic negotiation. I explain why Stalnaker’s traditional theory of conversation is ill-equipped to capture such communication about concepts, and outline an extension of his theory that can do better. In the following two sections, I then build on this framework, showing how it can be expanded to model communication of two more specific varieties of conceptual exploration. In section 3, I consider conceptual supposition and the counterconventional conditionals we use to communicate this reasoning. In section 4, I consider conceptual comparisons, and the metalinguistic comparatives we use to communicate these. In each case, I explain both how I propose to analyze the language in question, as well as how the form of reasoning being communicated is related to conceptual engineering — on the assumption that conceptual engineering is possible. In section 5, I’ll shift to consider how matters look from the point of view of someone with doubts about the feasibility of conceptual engineering (e.g., Cappelen 2018, Deutsch 2020). While suppositional and comparative reasoning about concepts are separable from conceptual engineering, their plausible cognitive connections with conceptual engineering provide abductive evidence against the engineering skeptic. Section 6 concludes.

2 Communicating about concepts

In this section, I will present a framework for modeling communication about concepts. Consider the case of Pluto. As the reader might recall, in 2006, the International Astronomical Union (IAU) revised the definition of planet. While Pluto had previously been classified as a planet, astronomers had become increasingly aware of the scientific inconvenience of this classification, given that more and more celestial objects were being discovered in the solar system that had equally good claims to planethood. The IAU thus redefined planet so as to require all planets to “clear their orbital neighborhood,” meaning they have to be substantially more massive than anything else near their orbit. Since Pluto’s orbit crosses Neptune’s, and Neptune is much larger than Pluto, Pluto is now classified as a dwarf planet, not a planet. Let’s consider an example of metalinguistic negotiation on this topic:
(1)  Ann:  Pluto is a planet.
Ben:  No it’s not, it doesn’t clear its orbital neighborhood.
Ann:  I don’t accept the IAU's definition! Pluto is a planet.

What we see here is a dispute about how to use the term *planet*, rather than a dispute about factual matters.¹

With her claim, *Pluto is a planet*, Ann isn’t communicating anything factual. Everything about Pluto’s orbit, size, etc., is common ground between the speakers here. Ann is instead advocating for a certain way of classifying things as planets. She is advocating for a certain understanding of the concept associated with *planet*, different from that adopted by the IAU.

Let me make a few notes about how I am understanding the key notions at work here. First, I take the meanings associated with terms to be concepts. (I leave open that concepts may also play other roles.) And I will take such concepts to determine an extension at a world. This means that if Ann and Ben agree about all the descriptive facts, but disagree about whether Pluto is a planet, then they must employ different concepts of planethood. Conceptual engineering, on this picture, involves changing which concepts we use, and more specifically changing which concepts are associated with which terms.² There is a difference, however, between what a speaker uses a word to mean on a given occasion — its *speaker meaning* — and what the word means in a more standing way in the linguistic community — its *semantic meaning* (Grice, 1989). To give a simple illustration, if I have a cat who behaves comically rather like a dog, I might jokingly say, *The dog wants to go for a walk again*. Everyone understands that by *the dog* I speaker-mean the cat and thus convey something true. But his doesn’t change the fact that the semantic meaning of *dog* includes only dogs and no cats.

Given the distinction between speaker meaning and semantic meaning, there are two different ways we could understand the project of conceptual engineering. First, it can involve changing, or aiming to change, which concepts speakers speaker-mean by their terms (Pinder, 2021). The scope of conversations that one aims to affect will depend on the case, though speakers usually will not be

---


²In understanding matters this way, I am broadly following Plunkett 2015. It is also similar to Cappelen 2018, though he prefers to avoid talk of “concepts” altogether, and simply focus on the meanings of terms. For an overview of recent approaches to conceptual engineering, see Cappelen and Plunkett 2020, Burgess et al. 2020.
content with changes that are too narrow or fleeting. Second, conceptual engineering can involve changing, or aiming to change, which concepts are the semantic meanings of terms. This is arguably a heavier lift, especially under the assumption that semantic meaning is often determined by factors external to the speaker, as on externalist metasemantic views in the tradition of Kripke (1980), Putnam (1973) and Burge (1979). Some, like Cappelen (2018, chap. 7) and Deutsch (2020), think that it is nearly impossible to effect such change in an intentional way. Koch (2021a,b), however, argues that speakers do have an important kind of control over semantic meaning. This control is collective and long-range, but that is just the same as with many other challenging projects that we nonetheless believe it worthwhile to participate in, like eliminating world hunger or mitigating climate change (cf. also Andow 2021). I’ll discuss skepticism about conceptual engineering more in section 5. For now, I will allow that conceptual engineering covers both changes in speaker meaning and semantic meaning.

Metalinguistic negotiation is often in the service of conceptual engineering understood in this way. One key way speakers advocate for a certain conceptual choice is by simply using words in the way called for by that choice. Ann uses planet so as to include Pluto in its extension as a way of advocating for such a concept of planethood. As I mentioned, I do not want to rule out that the semantic meanings of terms may be, at least in part, fixed externally. Even still, to accurately describe cases of metalinguistic negotiation, we should take speakers to be able to use words as if they are associated with concepts different from the standing semantic meanings of those terms. If enough speakers come to use terms in that way, it can result in successful conceptual engineering at the communal level, whether by influencing speaker meaning on a large scale or by eventually affecting semantic meaning. Note, however, that it’s also possible for conceptual engineering to take place without metalinguistic negotiation, e.g., through explicit discussion of concepts.

How should we theorize about the communicative effects of metalinguistic negotiation? I will next review Stalnaker’s theory of communication and explain why it is not up for the task. Then, I’ll present an extension of the theory that can do better.

On Stalnaker’s theory of communication, the common ground of a conversation — or the assumptions that are known to be mutually shared by all participants — is represented by a set of worlds. This is the set of worlds left open by those mutual assumptions. The content of an
assertion is also represented by a set of worlds: the worlds where the given assertion is true. The conversational effect of an accepted assertion is then to intersect the common ground with the content of the assertion — thus ruling out from the common ground all worlds where the assertion is false (e.g., Stalnaker 1970, 1999, 2002). Overall on this picture, communication is essentially information sharing. As speakers contribute to the conversation, the possibilities for how the world could be that are live for the purposes of that conversation get narrowed further and further. On this picture, speakers disagree when updating the common ground with both of their assertions results in the empty set, ruling out all possibilities.

Stalnaker’s theory is an elegant account of communication, and works very well for the progression of conversations involving the sharing of factual information. However, it faces a challenge from metalinguistic negotiation. The theory isn’t well-suited to capturing the conversational effect of claims, like those in (1), that do not communicate any new worldly information. Entering their conversation, we may assume that Ann and Ben share all factual commitments about Pluto (about its size, orbit, and so on). And we can furthermore assume that these commitments are all common ground between them. On Stalnaker’s theory, then, it’s not clear why Ann and Ben should be in disagreement. Their commitments about what the world is like are perfectly compatible. And yet, there is clearly some disagreement between them. The view of communication as information sharing has trouble explaining why.

Before moving on to an extension of Stalnaker’s theory that can deal better with metalinguistic negotiation, let me acknowledge a very natural thought that might arise at this point, and that might seem to avoid the need for any revisions. The natural thought is that Ann and Ben simply disagree about whether Pluto is a planet. To this, the answer is: Of course, yes, this is a way to describe their disagreement. But then the question is: Can Stalnaker’s theory tell a plausible story about what this disagreement consists in? The answer, I hold, is that it cannot. To see why let’s consider three options that may come to mind.

First, does the disagreement consist in disagreement about Pluto’s physical characteristics? No: We’ve already stipulated that those are all already common ground. Perhaps, then, the disagreement consists in disagreement about actual linguistic practices with the term planet? But again this doesn’t seem right. We could assume that Ann is under no misunderstanding about the
revision made by the IAU, and the extent to which that revision has been taken up or not by the wider linguistic community. We can assume this, too, is all common ground between our speakers. And yet a disagreement could persist.³

Finally, could Ann and Ben’s disagreement be about each others’ preferences concerning how to use the term *planet*? But these too could be common ground, and yet the disagreement persists. It’s not like Ben doesn’t know that Ann’s preferences about how to use the term differ from his. He isn’t disagreeing with her about her own preferences.

As we consider all these possible stories about what the disagreement between Ann and Ben boils down to, it becomes more and more apparent that they can engage in a sensible disagreement about whether Pluto is a planet, while not disagreeing about any worldly facts at all. They disagree about what concept to mean by *planet*. And this disagreement need not reduce to a factual disagreement. (Such disagreements might sometimes reduce to factual ones, but we don’t want our theory of communication to *require* that they always reduce in this way.) Stalnaker’s theory, in representing all content just with sets of worlds, doesn’t allow us to draw the intuitive distinction between a disagreement based in different factual commitments, and one based on different normative views about how to use concepts.⁴

To account for metalinguistic negotiations, we enrich Stalnaker’s theory so that speakers communicate not only information about the world, but also proposals for which concepts to use. More specifically, we take the common ground and the contents of assertions no longer to be represented by sets of worlds, but rather sets of *world-convention pairs* (Kocurek et al., 2020).⁵ A *convention* is

---

³This option also helps us see why Stalnaker’s (1978) diagonalization technique will not help in our case. The diagonal proposition expressed by Ann’s utterance of *Pluto is a planet* is that the referent of *Pluto* is in the extension of *planet*. This, on Stalnaker’s picture, is still a traditional possible worlds proposition, a function from worlds to truth values. So this approach would presume that the world settles the meaning of *planet*. However, we not want to presume this in the present discussion. And more importantly, making such a presumption would misrepresent what Ann and Ben take to be at issue between them. They are not disputing what *planet* means as determined by some worldly facts, but rather what it *should* mean — this is what makes it a metalinguistic negotiation.

⁴This is not so much intended as a criticism of the Stalnakerian framework, as much as a recognition that it was not yet designed to capture every aspect of communication. I am thus in agreement with Green (2017), when he remarks: “The [common ground]-context approach is [… ] highly abstract, so merely pointing out that it fails to account for an aspect of communication is an inconclusive criticism. Instead our question should be whether it can be extended or modified to account for such a phenomenon while preserving its spirit” (p.1589). Indeed, Stalnaker himself in more recent work has proposed extensions of his original framework that are in some ways similar to what I will propose here, except to account for communication about epistemic uncertainty rather than conceptual indecision. (See Stalnaker 2014, especially chapter 7.)

⁵Compare Barker 2002, which uses sets of world-delineation pairs to account for the conversational dynamics of vague predications; also MacFarlane 2016.
a function from linguistic expressions to concepts, and a concept, in turn, determines the extension of a term at a world. Thus, the convention that Ann adopts associates a different concept with the term *planet* than the IAU does. The content of Ann’s assertion of *Pluto is a planet* is the set of world-convention pairs according to which Pluto is in the extension of *planet* according to that convention at that world. World-convention pairs in this set will include, for instance: the pair of the actual world plus the folk definition of *planet* pre-2006, and the pair of a world where Pluto is much larger than it actually is and doesn’t cross Neptune’s orbit plus the IAU’s definition of *planet*.

Now, in the metalinguistic negotiation above, we assumed that the facts about Pluto’s size and orbit were already common ground. That is, the common ground already doesn’t contain any world-convention pairs where the world component is as described in the second pair just mentioned. With all facts about Pluto’s size, orbit and so on, common ground, Ann, in making

---

6Some discussions of metalinguistic negotiation, e.g., Plunkett and Sundell 2013, view it as entirely pragmatic, perhaps not requiring any revision to the view that assertive content is simply a set of worlds. On such an approach, what Ann and Ben are “really” disagreeing about — whether the concept of planethood *should* apply to Pluto — is not part of the literal content of their assertions (see also Sundell 2011, 2017). Questions remain for such a view, most importantly about the mechanism by which this additional, non-literal normative content is conveyed by means of what is literally expressed. Answers may be given here (e.g., Mankowitz 2021), and it is beyond my present aims to argue that they cannot succeed. However, such a pragmatic approach will not easily extend to the embedded uses of metalinguistic material that I will discuss in sections 3–4. (Similarly, Shan (2010) takes mixed quotation to operate at the level of semantics rather than pragmatics because of embedding constructions. Indeed, there are some commonalities between the present approach to metalinguistic constructions and Shan’s account of quotation, though he hews closer to the traditional Stalnakerian line than we do; see footnotes 3 and 8. Still, the relationship between quotation and metalinguistic uses merits further study.) The framework I develop here allows for a unified linguistic treatment of all kinds of communication about concepts, involving embedded and unembedded metalinguistic constructions. For approaches in a similar spirit, though not covering the same range of language-use, see Armstrong 2016, Muñoz 2019a,b.

7With contents as world-convention pairs, will we now overgenerate possible metalinguistic uses? An anonymous reviewer gives the following example: “I can’t be understood as making a metalinguistic proposal about conventions concerning the word *smartphone* when I say *My laptop is a smartphone* (even though my laptop has most of the features one expects of a smartphone — phone (VoIP), email, social media, etc.) — so it arguably should be an accessible reading.” In a suitable context, though, I believe that utterance could be understood as a proposal to change the concept of smartphone so it applies to (at least some) laptops. Imagine it, perhaps, as a response to friends who are bugging you to finally get a smartphone. Granted, the speaker would probably be taken to mean this as a bit of a joke; but the reason for that, I think, is that no one in the conversation seriously thinks the meaning of *smartphone* is up for grabs to such an extent that it might be modified to apply to laptops. A metalinguistic use, after all, is a proposal to change the conventions adopted in the conversation. For it to succeed, the change must be accepted by the interlocutors; and for it to be understood in the way the speaker intends, the interlocutors must realize that the relevant conventions are up for discussion. For this reason, I am not fully on board with Sterken’s (2020) claim that attempts to change language-use commonly cause miscommunication and confusion. While they no doubt occasionally do, it seems to me that, in usual cases, those in the conversation are aware, at least implicitly, that conceptual or linguistic matters are at issue. (This is still compatible with the linguistic interventions often being “disruptive” and “transformative”, as she holds.) Overall, then, I think there are no limits in principle on which sentences can be used metalinguistically. It’s just a question of how likely it is that the context will make such a use reasonable (cf., Kennedy and Willer 2016).
her assertion, is not making a proposal to narrow down the worldly possibilities in the common ground. Instead, she is making a proposal that, if accepted, would rule out certain conventions, or certain options for which concept to associate with the term  *planet*. On this picture, Ann and Ben do have a substantive disagreement. But it’s not because they can’t come to a shared view about what the world is like. Rather, it’s because they can’t come to a shared view about how to carve up that world with the concept of a planet. Accepting both Ann’s and Ben’s assertions would leave us with no world-convention pairs left in the common ground.

To sum up, enriching Stalnaker’s theory by replacing worlds with world-convention pairs allows us to make sense of metalinguistic negotiations as disputes about concept choices. In the next two sections, I will show how this framework can be further extended to capture speakers’ communication of their conceptual explorations.

3 Supposing alternative concepts

Let’s return to the metalinguistic negotiation from (1), but now add on a possible continuation:

(2)  

| Ann: | Pluto is a planet. |
| Ben: | No it’s not, it doesn’t clear its orbital neighborhood. |
| Ann: | I don’t accept the IAU’s definition! Pluto is a planet. |
| Ben: | If Pluto were a planet, there would be dozens of planets in the solar system. |

With his final assertion here (in bold), Ben is not advocating for a convention on which Pluto is a planet — quite the opposite. But he is communicating something about concepts. He is naturally taken to be defending the IAU’s definition of  *planet*, agreeing with their decision to have that term be associated with a concept that does not apply to Pluto, by pointing out an inconvenient consequence of a definition on which Pluto is a planet.

Sentences like the one Ben asserts last, repeated on its own in (3), have been termed COUNTER-CONVENTIONAL CONDITIONALS (Einheuser, 2006).

(3)  

If Pluto were a planet, there would be dozens of planets in the solar system.

In the dialogue between Ann and Ben, (3) is not naturally interpreted as a claim about what would happen if Pluto, say, cleared its orbital neighborhood. Rather, it is heard as a claim about
what would happen if we associated the term *planet* with a concept of planethood on which Pluto (with its actual characteristics) counts as a planet. Evaluating this conditional involves supposing hypothetical convention choices, rather than supposing hypothetical worldly possibilities.

Standard semantics for counterfactual conditionals are not equipped to predict this. On a Lewis-Stalnaker approach to counterfactuals, a conditional of the form *If A would B* is true just in case the closest worlds where *A* is true are also worlds where *B* is true. While this kind of analysis correctly predicts the meaning of many subjunctive conditionals, it doesn’t do well with counterconventionals, like Ben’s assertion of (3). Ben isn’t saying something about what would follow if the world were different in some way.

For Ben’s assertion to be correctly interpreted, as a claim about what would follow under alternative conventions, rather than under alternative worldly conditions, we can extend the conventionalist framework given in the previous section. The basic idea is that sentences are evaluated as true or false not at a world, but at a world-convention pair. Subjunctive conditionals in general can then shift not only the world used to evaluate the embedded material, but also the convention (Kocurek et al., 2020, Gaus, 2020).

On this view, which shares many basic features with the Lewis-Stalnaker approach, Ben’s assertion, *If Pluto were a planet, there would be dozens of planets*, is true just in case the closest world-convention pairs where *Pluto is a planet* is true are also pairs where *There are dozens of planets* is true. These closest world-convention pairs may differ from the original index in world, convention, or both. Which pairs are considered “closest” will depend on the context. This is just as in the traditional Lewis-Stalnaker semantics for counterfactuals. In the context of a metalinguistic negotiation, like (2), where what is at issue is clearly what concepts to use rather than what the world is like, it would be expected that the conditional should shift the conventional component more readily than the world component. Thus, with the addition of the convention

---

8One might think that shifting the world is sufficient, since the world can determine linguistic conventions. That is, one might adopt a diagonalization approach inspired by Stalnaker 1978. While this will work for some cases, it won’t capture the meaning of all counterconventionals. The reason is that we can shift to alternative conventions without shifting to a world at which those conventions are in effect, as shown in (i):

(i) *If Pluto were a planet, there would be dozens more planets in the solar system, even if no life had ever evolved.*

Here, the consequent is evaluated using shifted conventions for *planet*, but at a world where no life exists and so no linguistic conventions in effect. (Set aside the possibility of representational beings that are not alive.) See also Einheuser (2006) on “diagonal conventional conditionals” and “counterconventional conditionals”.

---
parameter in the semantics, we derive convention-shifting readings of conditionals as a result of the
general widely-recognized context-sensitivity of the selection function employed in the interpretation
of subjunctive conditionals.\footnote{For a formal implementation of this approach, see Kocurek et al. 2020, sec. 4.}

Counterconventional conditionals, on the present picture, thus allow us to communicate our
suppositional reasoning about alternative concepts. This suppositional reasoning is a form of con-
ceptual exploration. Next, let us ask: What is the relationship between such conceptual supposition
and conceptual engineering?

Counterconventional conditionals are used to reason suppositionally about what would follow if
we adopted different conventions. Conventions, i.e., the link between terms and concepts, are not
fixed once and for all. They can change over time, and our linguistic behavior has some effect on
these changes. As mentioned in section 2, we can arguably change both what concepts we speaker-
mean by terms, as well as (in less predictable and more long-term ways) the semantic meanings
of terms. However, we can not only advocate for different ways of using concepts, engaging in
conceptual engineering, but we can also reason about the implications of making such a change,
without yet committing to it. Conceptual supposition can thus help us see considerations for or
against certain concept choices. In this way, conceptual supposition can be in the service of projects
of conceptual engineering.

For example, in the dialogue above in (2), Ben is naturally taken to be pointing out an unde-
sirable consequence of counting Pluto as a planet. He is implicitly encouraging Ann to reason via
modus tollens that Pluto shouldn’t be classified as a planet, because classifying it as such would
have the inconvenient implication that there would be dozens more planets in the solar system. In
this case, then, he is advocating in favor of sticking with the IAU’s definition. Actively aiming to
preserve a certain way of using concepts is arguably as much a form of conceptual engineering as
actively trying to change how we use concepts (Lindauer, 2020). And so Ben’s use of the counter-
conventional conditional in (3) can be taken as a consideration relevant to conceptual engineering.

Counterconventional conditionals can of course also be used to endorse or encourage the con-
vention adopted in the antecedent. Consider (4), for instance.
(4) If oat milk were milk, then it could be sold in the dairy aisle.

We can imagine (4) being used by a dairy farmer to discourage the convention supposed in the antecedent, or equally by an animal rights activist to encourage it.

While it may be most common to use counterconventionals in arguing about the advisability of some convention choice, they need not only be used in this way. We can also use convention-shifting conditionals simply to consider what would follow if some convention were in force, without taking this to speak in favor of either adopting or resisting that convention. For example, someone might assert (5) to point out a surprising fact about the relative athletic prowess of horses and human beings. (Background fact: Secretariat was a renowned racehorse; see Ludlow 2014).

(5) If Secretariat were an athlete, then all the top athletes would be horses.

We could imagine this being a cooperative conversational move even if the speaker has no interest in engineering any of the concepts involved. They could be completely indifferent about whether racehorses count as athletes or not.

So we can see that, as activities, conceptual supposition is independent of conceptual engineering. It’s perfectly coherent to engage in the first without the second. Another way to see this is that suppositional reasoning about concepts doesn’t presuppose that conceptual engineering is even possible. Conceptual engineering requires some ability to change the concepts we use, whereas the kind of suppositional reasoning under consideration has no such requirement. Even if our concepts were fixed forever, we could still consider what would follow, were they different. Compare: I have no ability to change the outcome of the Second World War, but I can still consider and maybe even disagree with the hypothesis presented in The Man in the High Castle about what would have happened had the Axis won. We explore counterfactual possibilities all the time, and not only about things we have any power to change. So too, then, I can reason about counterconventional possibilities, even if I can’t in fact change what concepts are associated with which terms.

Despite this independence of conceptual supposition and conceptual engineering, conceptual supposition should not be ignored by theorists of conceptual engineering. This is not only because of its role towards figuring out which conceptual changes to endorse. It is also because, as I’ll discuss in section 5, our capacity for suppositional reasoning about concepts (as well as other
forms of conceptual exploration) provides abductive evidence against skepticism about conceptual engineering.

4 Comparing concept choices

Let’s again return to our metalinguistic negotiation, but this time consider a different rejoinder from Ben.

(6) **Ann:** Pluto is a planet.
    **Ben:** No it’s not, it doesn’t clear its orbital neighborhood.
    **Ann:** I don’t accept the IAU’s definition! Pluto is a planet.
    **Ben:** Come on, Pluto is more an asteroid than a planet.

This final comment from Ben (in bold) is fairly reasonable. The trans-Neptunian objects, of which Pluto is the most famous, form the Kuiper belt, which is rather similar to the asteroid belt between Mars and Jupiter. In fact, Ceres, the largest object in the asteroid belt, was considered a planet for many years, before increasing discoveries of similar bodies started to make that inconvenient. Now, Pluto is classified as a dwarf planet, and neither an asteroid nor a planet. But Ben’s claim that between the latter two choices, it is more aptly characterized as an asteroid than a planet, makes some sense.

With this final assertion, then, Ben certainly isn’t advocating for a convention on which Pluto is a planet. He also isn’t necessarily advocating for a convention on which Pluto is an asteroid. But he is communicating something about his relative commitments to choices regarding the two concepts.

In (6), Ben communicates something about his conceptual commitments with what has been called a metalinguistic comparative (repeated in (7) on its own).10

(7) Pluto is more an asteroid than a planet.

Metalinguistic comparatives are distinct from “ordinary” comparatives, which compare the degrees to which individuals possess gradable properties, as in (8).

---

(8) Alice is taller than Sam.
≈ The degree to which Alice is tall is greater than the degree to which Sam is tall.

By contrast, the metalinguistic comparative in (7) makes sense, even if we assume that being an asteroid and being a planet are all-or-nothing matters. Instead, paraphrases for the metalinguistic comparative tend to appeal to comparisons of appropriateness of different ways of talking, as in (9).

(9) Pluto is more an asteroid than a planet.
≈ It’s better to say that Pluto is an asteroid than to say that it’s a planet.
≈ It’s more appropriate to categorize Pluto as an asteroid than as a planet.

In fact, in some languages, such as Japanese, metalinguistic comparatives are expressed using morphemes that also mean to say (Sawada 2010, Morzycki 2011).

We can use metalinguistic comparatives to communicate our relative commitment to different convention choices. For instance, with (7), Ben communicates that he is more committed to a convention on which Pluto is an asteroid compared to one on which Pluto is a planet. In this section, I will propose an approach to metalinguistic comparatives that makes good on this idea, building on the framework from sections 2–3.

In order to see the motivation for the proposal, let’s begin by considering: If metalinguistic comparatives compare what is, in some sense, better or more appropriate to say, what kind of appropriateness is at work here? A very general notion will not work. This is forcefully illustrated by Morzycki (2011, p. 47), with the following example.

(10) [Herman approaches the bereaved at a funeral and says, “Sorry your mother croaked.”
Herman’s friend takes him aside and says . . .]

a. It’s more appropriate to say “She passed away” than to say “She croaked.”
b. ??She more passed away than croaked.

The idea that metalinguistic comparatives communicate some kind of comparative appropriateness is on the right track. It’s just that it’s not appropriateness, all things considered. Rather, they communicate a specifically semantic kind of appropriateness. This can be captured by extending the conventionalist framework laid out above to account for metalinguistic negotiations and countercon-
ventional conditionals. There, we just took speakers to have commitments to certain conventions, in addition to their commitments to certain ways the world could be. Now, we must complicate the picture to recognize speakers as having a commitment ordering on possible conventions.

Speakers don’t just adopt certain conventions and reject others. They also have more nuanced attitudes: Among all the conventions they don’t adopt, for instance, they view some as better contenders than others. This is analogous to the relationship between belief and credence. There are some propositions an agent believes and others that they fail to believe. This, however, is not exhaustive as a description of an agent’s doxastic state. Agents also have varying credences in propositions. Though I may neither believe that it’s currently raining in Montreal, nor that it’s currently sunny there, I have a higher credence that it’s raining. Representing doxastic states using credence and not merely belief is essential for understanding behavior under uncertainty, but also, according to many theorists, for modeling probabilistic language, as in (11) (Moss, 2018, Yalcin, 2010).

(11) It’s more likely that it’s raining in Montreal than that it’s sunny there.

With (11), the speaker expresses that their credence that it’s raining in Montreal is higher than their credence that it’s sunny there. They express their comparative credences in the two propositions.

Wellwood (2014, chap. 6) in fact proposes that metalinguistic comparatives be analyzed as claims about comparative credences.\textsuperscript{11} This is not quite right, however, given that metalinguistic comparatives and claims of comparative likelihood are not equivalent:

(12) Pluto is more an asteroid than a planet.

\[ \not\approx \text{It’s more likely that Pluto is an asteroid than a planet.} \]

We may, for instance, be completely certain that Pluto is neither an asteroid nor a planet, but still accept the metalinguistic comparative. But though an analysis of metalinguistic comparatives in terms of comparative credences is inadequate, the present proposal will be importantly analogous to it: We must simply replace comparative credences with comparative commitments to conventions.

\textsuperscript{11}She calls them “categorizing comparatives”, which is an apt label on my view; still, I stick with the more standard terminology of “metalinguistic”. Indeed, one might view the conventionalist framework as building into the semantics many things that have traditionally been viewed as “metalinguistic”. Note that she later moved away from the credence view in Wellwood 2019.
To account for metalinguistic comparatives, then, we should supplement the conventionalist framework given in the previous sections as follows. We model speakers as not just accepting or rejecting certain world-convention pairs, but ranking them. Sentences are then evaluated not only relative a world and a convention (as in section 3), but also relative to an ordering of world-convention pairs. The metalinguistic comparative in (7) is then true relative to such an ordering just in case, holding fixed the world, there is a convention making \textit{Pluto is an asteroid} true that ranks higher than any convention making \textit{Pluto is a planet} true. In asserting this metalinguistic comparative, Ben expresses his stronger commitment to a convention on which Pluto is an asteroid than to one on which Pluto is a planet.\footnote{For a formal implementation as well as more discussion of how this approach compares to previous analyses of metalinguistic comparatives, see Rudolph and Kocurek 2020.}

Recall that we are taking conventions to assign concepts to terms. Thus, on the proposed analysis, metalinguistic comparatives ultimately serve to express our comparative commitments to certain conceptual choices. Ben, in the dialogue above, is more committed to a concept of asteroid on which Pluto counts as an asteroid than he is to a concept of planet on which Pluto counts as a planet. This kind of comparison is another form of conceptual exploration. Next, let us consider the relationship between conceptual comparison and conceptual engineering.

In many cases, comparing concept choices is in the service of projects of conceptual engineering. In the dialogue in (6), we would assume that Ben would sooner advocate for changing the concept associated with \textit{asteroid} so as to include Pluto than for changing the concept associated with \textit{planet} to do the same. At the least, he is certainly advocating \textit{against} a certain view about what concept to associate with \textit{planet} — one that differs from Ann’s.

However, metalinguistic comparatives can also be used simply to express comparative commitments to concept choices, even if one has no interest in engineering the concepts involved. For instance, it’s perfectly coherent for someone to assert (13).

\begin{equation}
\text{(13) We all agree that Pluto is a dwarf planet, so neither an asteroid nor a planet. Still, it’s more an asteroid than a planet.}
\end{equation}

Conceptual comparison is also independent of conceptual engineering in that it (like conceptual supposition) doesn’t require that we have any ability to change how we use concepts in the way
that conceptual engineering itself does. Compare: I can compare my relative preferences in certain worldly outcomes — say, having a winning as opposed to a losing lottery ticket — even if I have no control over whether I get what I prefer. Similarly, I can compare my relative commitments to different concept choices even if I can’t have any effect on which are ultimately adopted. Nonetheless, as I’ll discuss more in section 5, our capacity for conceptual comparison holds lessons for theorists of conceptual engineering.

5 Lessons for conceptual engineering

In sections 3 and 4, I discussed how two forms of conceptual exploration — conceptual supposition and conceptual comparison — are often used in the service of conceptual engineering, and yet are also independent of it. They are independent of conceptual engineering in that even if we had no control over the concepts associated with our terms in unembedded contexts, we could still reason under the supposition of alternative concepts, as expressed with counterconventional conditionals; and we could still compare possible conceptual choices, as expressed with metalinguistic comparatives. One upshot of this situation is that even theorists who are doubtful about the prospects of conceptual engineering cannot ignore the role of alternative concepts in theorizing about human cognition and communication. Whatever control we may or may not have over the concepts associated with our terms, it’s an empirical fact that we communicate meaningfully about possible alternatives with certain forms of language. Semantic theory must incorporate shifts in concepts in some form, if it’s to adequately account for counterconventional conditionals, metalinguistic comparatives, and likely other constructions beyond these.\(^\text{13}\)

Furthermore, conceptual exploration undermines skeptical views about conceptual engineering. Assume that the semantic meaning of at least many of our terms is determined in part by factors external to speakers and their mental states. This is commonly accepted, following Kripke (1980), Burge (1979) and Putnam (1973). I can’t just make the term water refer to whatever I want. Given the circumstances in which the term came into use and continues to be used, and the fact that the stuff in lakes and rivers and coming out of taps around here is \(\text{H}_2\text{O}\), that is what water refers to — whether people know it or not, and whether people like it or not. Though meaning change is

\(^{13}\)See, for example, Muñoz 2019a, Kocurek et al. 2020 on attitude reports and Armstrong 2013, chap. 3 on loose speech.
possible within externalist metasemantics, theorists like Cappelen (2018) and Deutsch (2020) argue that the way this meaning change takes place leaves the would-be conceptual engineer with a serious implementation problem. The reason is that the determinants of meaning are inscrutable to us, and even if we could find out about them, they are largely outside of our control (Cappelen, 2018, chap. 7). Deutsch (2020, p. 3953) further emphasizes that the problem isn’t due to externalism per se, but will arise with any metasemantic theory that takes the determinants of semantic meaning to be suitably disconnected from speakers’ intentions to start using terms in new ways.

As I mentioned in section 2, there are defenses of conceptual engineering against this challenge. Pinder (2021) defends conceptual engineering, while conceding that it only targets speaker meaning and not semantic meaning. Along similar lines, Flocke (2020) takes conceptual engineering to change how speakers assess the truth of propositions, which she takes to be compatible with no change in the semantic meanings of the terms involved. On the other hand, Koch (2021a,b) defends conceptual engineering even as targeting semantic meanings. Though we lack individual immediate control over semantic meanings, we nonetheless can exercise a kind of long-range collective control over them. And this is enough, he holds, to make conceptual engineering worthwhile to engage in. There is no guarantee of success, but the same goes for many long-term collective projects we nonetheless reasonably choose to be a part of, like combating climate change. These styles of response also naturally complement one another. The way to bring about long-range collective semantic change, of the sort Koch believes in, is likely in large part by influencing what speakers speaker-mean by their terms, or how they evaluate the truth of propositions involving them.

These defenses of conceptual engineering are compelling. And we can bolster them further by bringing conceptual exploration into the picture. As we’ve seen, conceptual exploration is often in the service of conceptual engineering. Thus, we might think that if we never had the power to engineer our concepts, it would be unexpected for us to have the capacities for conceptual exploration discussed above.

---

14 Note that Cappelen doesn’t brand himself a full engineering skeptic. He holds that, even in the face of the challenges he identifies, “we will and should keep trying” to engage in conceptual engineering (p. 72).

15 Note that even a view of conceptual engineering that takes it solely to operate on speaker meaning doesn’t make such changes trivial or necessarily easy to accomplish, as Fischer (2020) argues drawing on cognitive linguistic evidence.

16 For an overview of possible responses to the implementation problem, see Jorem 2021.
Consider conceptual supposition first. As I’ve mentioned, it is certainly possible to reason suppositionally about possibilities — both conceptual and factual — that are outside of our control. Still, we might think that if we never had the power to engineer our concepts, it would be unexpected for us to have the capacity for suppositional reasoning about concepts. Compare: If we never had any power to change the course of events, our capacity for counterfactual reasoning would be difficult to understand. Humans’ ability to imagine and reason about alternative possibilities is arguably closely tied to our ability to make effective plans and decisions. It’s a way to, in a sense, experiment about how to live without actually having to live through the experiments oneself (e.g., Byrne 2005, 2016; Starr 2019, sec. 1.2). In a similar way, then, we might think that our capacity for counterconventional reasoning would be hard to understand if we never had the ability to engineer our concepts. Just as counterfactual reasoning allows us to experiment with alternative courses of events without living through them, counterconventional reasoning allows us to experiment with concept change without adopting such change.

Next, consider conceptual comparison. Again, it is certainly possible to compare our commitments to options that are outside of our control to select between. Still, we might think that if we never had the ability to engineer concepts, it would be unexpected for us to have the capacity to make these kinds of conceptual comparisons. Compare: If we never had control over outcomes, would it make sense to reason about our preferences among them? Maybe it would; but one could also perhaps argue that the point of such preferences is ultimately to direct us to bring about outcomes that are higher in our ranking rather than lower. If that is right about the point of relative preferences in worldly outcomes, then something similar could hold for relative commitments to conceptual choices. Just as preferences about the world guide us in bringing about better factual outcomes rather than worse, so too commitment orderings on concept choices can guide us to do the same when it comes to changes in the concepts associated with terms.

Moreover, if we don’t have the power to engineer our concepts, then much (though not all) of what we aim to do in conceptual exploration rests on a mistake. Recall the dialogue between Ann and Ben about whether Pluto is a planet. In the version from section 3, Ben utters the counterconventional conditional, repeated in (14).
(14) If Pluto were a planet, there would be dozens of planets in the solar system.

Sometimes people may utter counterconventional conditionals simply to make observations about alternative concept choices. But in the context given above, Ben is naturally interpreted as doing more: as voicing support for the International Astronomical Union’s conceptual choice that resulted in Pluto no longer counting as a planet. Is Ben operating under the mere illusion that anyone has control over what concept to associate with *planet*? It would be odd to resort to this view, given the plausible accounts of conceptual engineering mentioned above. One might claim that *planet* is a special case, given the authority of the IAU over astronomical definitions. But there are equally natural uses of counterconventional conditionals, in the apparent service of conceptual engineering, where no such official authority structure is in place. For instance, one could imagine the activist mentioned in the introduction using (15) in advocating for a more systemic definition of *racism*.

(15) If only thoughts and feelings could be racist, then it would be harder to address some of the biggest challenges facing Black Americans.

This doesn’t have to be directed at dictionary editors for it to have a clear aim. We can easily imagine it said among a group of peers negotiating how to talk about racism.

Similarly, recall the alternative conversational ending from section 4, with Ben uttering the metalinguistic comparative repeated in (16).

(16) Pluto is more an asteroid than a planet.

Again, sometimes metalinguistic comparatives are used outside of the context of conceptual engineering. But in the scenario we considered, Ben is naturally heard to be advocating against a return to the folk concept of planet — something that his interlocutor, Ann, supports. If this matter were outside of any speakers’ intentional control, this would be a hard conversational move to comprehend.

Sometimes, of course, people engage in practices that rest on mistakes (human sacrifice to appease deities, say). However, given the availability of promising accounts of how conceptual engineering works, I see no reason to resort to such an error theory about conceptual exploration. The cognitive capacities involved in conceptual exploration, as well as its common rationale, all
provide abductive evidence against skepticism about conceptual engineering.

6 Conclusion

Philosophers have given a lot of attention to conceptual engineering, but less attention to related activities that also reflect our ability to examine the concepts we use. We should recognize and theorize about conceptual exploration, specific examples of which include supposing alternatives concepts and comparing concept choices; and there are surely other modes of conceptual exploration left to be explored.

These activities are unified psychologically in that they all involve reasoning about our own conceptual machinery. Our capacity to engage in these forms of reasoning also emerges in particular kinds of language use. Adequate analyses of metalinguistic negotiation, counterconventional conditionals, and metalinguistic comparatives must capture the fact that, with them, we communicate not only about worldly possibilities, but also about the concepts that we use to categorize things in the world. Conceptual exploration is thus also unified linguistically, in that the language of conceptual exploration calls for us to enrich our linguistic theory in a way that incorporates conventions. The framework that I’ve introduced and expanded throughout this paper is a unified way of doing this, building on more traditional theories of meaning and communication that were only designed to capture communication about factual matters.

Additionally, recognizing our capacity for conceptual exploration undermines skeptical views about conceptual engineering. Anyone who holds that conceptual choices are largely outside of our intentional control has to take speakers to be operating under a serious misconception in much of their conceptual exploration. They also have to contend with the fact our very possession of capacities for conceptual exploration makes most sense under the assumption that conceptual engineering is possible. When paired with extant accounts of the workings of conceptual engineering, the present examination of conceptual exploration bolsters the view that conceptual engineering is a genuine practice that ordinary speakers as well as philosophers engage in — granted with varying levels of self-consciousness and success.
Acknowledgments

For helpful feedback, I wish to thank audiences at the 2021 Eastern APA and the Auburn University Philosophical Society. Special thanks to Kelly Gaus, Arc Kocurek, and an anonymous reviewer for this journal.

References


