Inscrutability and Its Discontents

Laura Schroeter and François Schroeter

University of Melbourne, Philosophy Discipline, Melbourne, Australia
Email: laura.schroeter@unimelb.edu.au; fschro@unimelb.edu.au

Abstract

Our main focus in this paper is Herman Cappelen's claim, defended in *Fixing Language*, that reference is radically inscrutable. We argue that Cappelen's inscrutability thesis should be rejected. We also highlight how rejecting inscrutability undermines Cappelen's most radical conclusions about conceptual engineering. In addition, we raise a worry about his positive account of topic continuity through inquiry and debate.

Keywords: Cappelen; semantic externalism; metasemantics; conceptual engineering

We are sympathetic to many of the core theses of Herman Cappelen's *Fixing Language* (2018). In particular, we think Cappelen is right to emphasize the importance of metasemantics—a theory of what makes it the case that expressions have the meaning they have—for a philosophical approach to conceptual engineering.[1] And, like Cappelen, we favor a metasemantics in the externalist tradition of Kripke (1980), Putnam (1975), Burge (1979, 1986), and Williamson (1994, 2007). Cappelen emphasizes three points about this tradition, which we think are correct (63):

i.   *External facts can affect reference*: Facts about individuals' physical, social, and historical environment can be relevant to determining the extension and intension of their expressions.
ii. *Possibility of "massive, fundamental" error*: Both individuals and their communities can have false assumptions (or unsatisfied desires) about (a) central or essential characteristics of the reference, and (b) whether there is any reference at all.
iii. *No "common creed"*: Semantic coordination among different speakers using a term does require that they all accept "a common creed"—a core set of assumptions or attitudes associated with an expression. *A fortiori*, semantic coordination (competence with the same meaning, coordination on the same topic) cannot be explained in terms of matching patterns in individual speakers' understanding.

Unlike Cappelen, we are not skeptics about concepts, but this difference will not play a significant role in this paper. Our main disagreement with Cappelen concerns his claim that reference is radically inscrutable. As we'll see in the next section, Cappelen takes semantic externalism to entail that the facts that determine reference and reference change are too complex

---

[1] "At the foundation of a theory of conceptual engineering is a theory of metasemantics" (7). Unless otherwise noted, all page references are to Cappelen (2018).

for us to fully grasp (72–74). As a result, in most cases, we cannot know the reference of our terms at a time, or whether their reference changes over time or between individuals.

In this paper, our main focus will be Cappelen's Inscrutability thesis. We'll argue that it should be rejected. We'll also highlight how rejecting inscrutability undermines Cappelen's most radical conclusions about conceptual engineering. In addition, we'll raise a worry for his positive account of topic continuity through inquiry and debate.

1. Semantic externalism and inscrutability

Cappelen doesn't argue for semantic externalism, nor does he propose a fully worked-out externalist metasemantics. Nonetheless, he takes his argument for inscrutability and lack of control over reference to follow from "moderate versions of externalism" in the tradition of Putnam, Kripke, Burge, and Williamson (9). What he takes from these theorists is that a range of external factors can play a role in helping to fix reference, and this fact undermines the possibility of knowledge and control over reference:[2]

> According to externalists, the meaning of our words can be influenced by features of the past, including introductions of expressions (such as pointing and stipulations, on the understanding that these can be massively messy), and communicative chains (the "passing along" of expressions where this is accompanied by something like reference-preserving intentions). Sources of information in the past—of the kind Evans (1973) talks about—can also be relevant to fixing reference. Other people can have an effect on our extensions: both Burge and Putnam provided good evidence that people classified as experts can play an important role. The total interaction between speakers' use and dispositions can also play an important role.
> This bundle of metasemantic claims entails both an epistemic point—that the metasemantics of our natural language terms are inscrutable—and a metaphysical point—that we have no control over the metasemantics. (73)

We'll focus here on Cappelen's inscrutability claim. What he wants to establish is that "[t]he process governing particular changes [in reference] is typically incomprehensible and inscrutable" (53). This claim plays a central role in Cappelen's understanding of conceptual engineering. Because we can't know which changes in our linguistic practices would trigger a change in reference, any effort to systematically plan and implement changes in the reference of our words is bound to be highly unreliable.

> The fact that conceptual engineering is inscrutable and out of our control means that it is also possible (sometimes I think even likely) that those who try to achieve good ends through conceptual engineering will end up causing harms they didn't intend. We have no prima face reason to think the process is typically one that leads to amelioration rather than degeneration. (159)

Cappelen's picture of conceptual engineering is thus a matter of flailing about in the hopes that one will stumble in a better direction. But we can't know whether we have satisfied our wishes, and we have an equally good chance of stumbling in the wrong direction.

So how exactly does moderate externalism generate these surprising—and alarming—epistemic consequences? Cappelen argues from inscrutability of the metasemantic facts to the inscrutability of semantic facts about reference and referential change. We can think of the metasemantics as an "interpretation function," which takes as input all the potentially relevant empirical facts about the

---

[2]Cappelen moves freely between talk about reference, talk about extension or intension, and talk about semantic values. We'll follow this practice here.

use of a term and delivers as output the reference of that term (its extension and intension) on that occasion of use. In effect, Cappelen offers two reasons why moderate externalist metasemantics makes facts about reference and referential change inscrutable. First, we cannot know all the relevant empirical inputs into the interpretation function:

> […] to figure out the current intension of a term, you would need information about the past, about introductory events, and communicative chains. It is indisputable that we don't have this information and never will. (73–74)

Second, Cappelen stresses at many points that the interpretation function itself is too "messy" for us to have reliable access to semantic facts about reference and referential change.

> In most cases the detailed mechanisms that underpin particular instances of conceptual engineering are too complex, messy, nonsystematic, amorphous, and unstable for us to fully grasp or understand. (72)
> [T]he actions and intentions of groups have at best a messy and unpredictable effect on our semantic values. (74)

This second worry, Cappelen argues, shows that reference will be just as inscrutable for an internalist as for an externalist: even if all the inputs into interpretation function are internal states, the interpretation function itself will be too messy and nonsystematic to afford access to reference and referential change (81–83).

To sum up, Cappelen claims that moderate externalism makes reference inscrutable because:

 (i)   The *inputs into interpretation* are not (and cannot be) fully known.
(ii)   The *interpretation function* is not (and cannot be) fully known.

Before we evaluate the strength of Cappelen's argument, we need to first clarify the intended conclusion. Presumably, inscrutability has something to do with what can be known given our limited access to empirical facts and our finite cognitive powers. But what exactly does it mean to say a term's reference (or a change in its reference) is inscrutable? Cappelen doesn't say. Here are two possible glosses on what we cannot know about intensions, which represent two ends of a spectrum:

(a) *Unknown Boundaries*: Given our actual cognitive limitations, we cannot know how to draw a precise boundary between the extension and antiextension of our term (or its extension/ antiextension for possible worlds). As a consequence, we cannot know when its extension has changed over time.
(b) *Cluelessness*: Given our actual cognitive limitations, none of our assumptions about what's represented by our words can be known to be true. In particular, our methodological assumptions about how to get closer to the truth about what's represented cannot be trusted to be reliable.

Cappelen's larger dialectic may suggest he is interested in (b) rather than (a), the weaker thesis. After all, Cluelessness seems to fit with Cappelen's view that we are subject to "massive, fundamental" errors about the reference of our own terms, and that our efforts to change the reference may lead us in the opposite direction to the one intended. In contrast, the lack of knowledge of where exactly to locate the precise boundary between the extension and antiextension for terms like "bald" or "marriage" does not seem to threaten our ability to eliminate "massive and fundamental errors" about the nature of these things. Nor does our inability to know precisely when a change in extension has occurred threaten our ability to foresee how broad changes in use could eventually result in a referential shift that better meets our interests. So on the face of it, (b), the stronger conclusion, seems required for Cappelen's ultimate conclusions about the radical inaccessibility of

semantic facts and our powerlessness to affect them through planning. However, Cappelen himself seems to deny that we are entirely clueless about reference and change reference of our terms. His point, he suggests, is that metasemantic theory is hard and theorists and ordinary speakers alike can have difficulties in deciding particular cases (66).[3] In that case, his conclusion should be something like (a). This interpretation, however, significantly deflates the inscrutability of reference and seems to undercut Cappelen's worries about planned conceptual engineering.

In evaluating Cappelen's argument, we propose to remain neutral between (a) and (b). Our questions here are: Does either conclusion follow from Cappelen's metasemantic premises (i) and (ii)? And is either conclusion a consequence of moderate externalism in the Putnam-Kripke-Burge-Williamson tradition?

Let's start with conclusion (b), Cluelessness. Has Cappelen given us any reason to think that we are clueless about the reference of our terms? Here, we think the answer is clearly no. Not only does Cappelen's metasemantic argument fail to support this conclusion, but there are strong reasons why moderate externalists do and should reject this claim.

It's important to be clear just how radical Cluelessness is. On this view, our understanding could be so impoverished that all beliefs about reference could be false, and it would be impossible to improve that understanding through rational inquiry. This is a *much* stronger thesis than the possibility of "massive, fundamental" error that we agree is a central commitment of moderate externalism. Here is our gloss of the latter thesis:

> *Possibility of "massive, fundamental" error*: Both individuals and their communities can have false assumptions (or unsatisfied desires) about (i) central or essential characteristics of the reference, and (ii) whether there is any reference at all.

Notice that this is a claim about the fallibility of specific assumptions *considered individually*. According to moderate externalists, virtually any specific assumption about the reference of our term could, in principle, turn out to be false—even if that assumption strikes us as a central or necessary truth. Cats could turn out to be robots; it could turn out that there is no water; it could turn out that Donald Trump is very intelligent. In contrast, Cluelessness entails that the *totality* of our understanding of a term could be radically off base. Perhaps our use of "3" really does refer to Julius Caesar; we are in no position to know. Clearly, it would be a mistake to take the mere possibility of error to support Cluelessness: even if any of our assumptions could turn out to be wrong, it does not follow that the all our assumptions could be wrong at once.

Moreover, Cluelessness itself is a highly implausible view. The problem is not just that it flies in the face of interpretive charity. The real problem is that it's hard to see what the theoretical point could be of assigning semantic contents to terms if those contents are completely unconstrained by speakers' own understanding and history of use. There are basically two types of theoretical role that semantic contents play:

(a) *Explanatory roles*: e.g., predicting and explaining individual speakers' behavior, or explaining the propagation of a term within a community
(b) *Normative roles*: e.g., explaining which uses of a term are rational, semantically correct, or true

---

[3]"What I do assume is that reference shifts happen. That understanding the exact underlying mechanism(s) that trigger reference shift is hard is itself an important data point for my theory. The experts don't know how it happens, much less do ordinary speakers. *It is not that we have no clue*: Evans has suggested we focus on dominant source of information, Devitt has proposed a related view and Kripke proposes […] that under certain conditions the intention to refer to a particular object overrides the intension to preserve reference. Maybe one of these or some combination of them is along the right lines—but even so, it provides precious little guidance in particular cases" (66; emphasis added).

Cluelessness entails that the correct semantic assignment could be entirely divorced from the beliefs, history of use, and implicit reflective dispositions of speakers and their linguistic communities. But in that case, correct semantic assignments would be useless: we couldn't use them in explanations, and they would set irrelevant normative standards. This conception of reference seems utterly pointless.

Clearly, this radical metasemantic thesis is not a corollary of moderate versions of externalism. The original externalist arguments of Putnam, Kripke, and Burge all presuppose that our reflective, empirically informed judgments about the reference of our terms are a reliable guide to the semantic facts. Most metasemantic theorists—internalist and externalist alike—take reliability to be a constraint on an adequate metasemantic theory.[4]

Nor does Cluelessness follow from Cappelen's two metasemantic premises:

  (i) The *interpretation function* is not (and cannot be) fully known.
  (ii) The *inputs into interpretation* are not (and cannot be) fully known.

Even if we don't know the *precise* interpretation function and we don't know *all* of the relevant empirical inputs, this does not establish that we could be entirely in the dark about the reference of our terms.[5]

In sum, the metasemantic argument does not support Cluelessness, and moderate externalists have every reason to reject that conclusion. So this version of the argument provides no support for Cappelen's suggestion that we are condemned to flail in the dark when we engage in conceptual engineering.

Let's now consider whether the argument supports the weaker conclusion (a), Unknown Boundaries. If we assume that the intensions of ordinary terms like "bald" or "marriage" are perfectly precise, how can we possibly discover where this sharp boundary falls? Which number of hairs marks the precise cutoff point for baldness? Given the assumption that the interpretation function determines precise boundaries for these terms, it strikes us as plausible that figuring out exactly where the sharp boundary falls would require (i) knowledge of the precise interpretation function, and (ii) full information about all the potentially relevant empirical inputs needed to fix the reference of "bald" as, say, ≤1,562 hairs. And these two facts, we agree, are something we are not generally in a position to know. So we agree the argument would be sound if it's true that the extensions of ordinary terms have perfectly sharp boundaries.

However, the existence of sharp boundaries is not entailed by "moderate versions of externalism." Most theorists deny that the interpretation function yields sharp boundaries for ordinary terms like "bald" or "marriage." Indeed, it's widely agreed that virtually any term in ordinary language is vague to some degree.[6] If there are no sharp boundaries for intensions and extensions of

---

[4]Proponents of such a constraint include Lewis (1974), Peacocke (1992), Jackson (1998), Chalmers (2006), Yablo (2008), Schroeter and Schroeter (2015). Williamson (1994) agrees that our reflective judgments allow us to identify the extension within a reasonable margin of error. Ruth Millikan's teleosemantics supports perhaps the most radical form of epistemic unreliability of any contemporary theorist (1993). But even she agrees that our *past* uses of a term must have been associated with a sufficiently accurate understanding of their reference to be favored by natural selection. So Cluelessness *over time* is impossible on her view.

[5]Indeed, there's a pragmatic problem in proposing *any* positive argument for Cluelessness. If Cluelessness is true, there is no reason to suppose we have any special insight into the reference of "reference." So why should we accept Cluelessness as a correct thesis about the nature of reference? For all we know, traditional descriptivism has just as good a chance of being true. Or it could turn out that reference relation is actually constituted by immaterial links of magical fairy floss.

[6]Williamson (1994) is the exception. His epistemic view of vagueness is motivated by a desire to preserve the law of excluded middle. But Williamson's view of vagueness is highly controversial and cannot be read into the commitments of "moderate versions of externalism." The issue of vagueness played no role in the classical externalist arguments of Putnam, Kripke, and Burge. And prominent externalists such as Millikan (1984, 2017) and Boyd (1980, 2013) explicitly embrace the idea that there is significant vagueness in the reference determination relation and the resulting referential assignments.

ordinary empirical terms, then the fact that such boundaries would be epistemically inscrutable is a moot point.

But perhaps moderate externalists are committed to Unknowable Boundaries independently of any interest in eliminating vagueness. Following Dorr and Hawthorne (2014), Cappelen suggests that the extension of a term like "marriage" could be so sensitive to slight changes in its use that its extension (*together* with its vague penumbra) could shift during the course of a single conversation (66) or even within a single second (111). If the interpretation function is exquisitely sensitive in this way to minor shifts in empirical facts, then perhaps the only way to know about a tiny change in extension might be to have (i) a correct theory of the interpretation function and (ii) knowledge of the totality of the empirical inputs for a particular use of a term. But as Cappelen argues, such knowledge is beyond our cognitive powers. So if Dorr and Hawthorne are right about the extreme sensitivity of the interpretation function, it's plausible Cappelen's argument for (a) is sound. If tiny shifts in usage are apt to generate tiny shifts in reference, then we are not in a position to know about these fine-grained referential changes.

However, this is not moderate externalism. Even Dorr and Hawthorne admit that positing constant tiny shifts in reference is a radical idea (2014, 284). Moderate externalists certainly agree that reference can shift gradually over time without our realizing it, as in Burge's (1988) slow-switching case and Evans's (1973) Madagascar case. But these shifts depend on protracted and widespread changes in the patterns of use within a community. Take the case of the word "corn." Originally it was used to refer to any type of grain, but in North America the term gradually came to be used exclusively to talk about the local grain—maize. This shift in patterns of use eventually changed the term's reference in the American dialect. The general empirical and semantic facts grounding this shift are well known. Of course, there can be penumbral cases where the reference of "corn" is indeterminate; and there can be cases where speakers make mistakes about whether two uses of the term are coreferential. However, moderate externalism holds that speakers' assumptions about sameness of reference are normally highly reliable and epistemically warranted.

This is not the place to engage in a systematic evaluation of Dorr and Hawthorne's arguments for extreme instability of reference. But given the key role their view seems to play in Cappelen's thinking, we'd like to make two quick points.[7]

First, a broad philosophical point: The point at issue for Cappelen is whether ordinary speakers have reliable and epistemically warranted access to facts about the reference and coreference of their terms. An acceptable answer to this question, we submit, must cohere with the core theoretical roles played by semantic contents. If radically unstable and inscrutable extensions cannot play the core theoretical roles we associate with semantic contents, then any metasemantic theory that assigns such extensions must be false. On the face of it, referential assignments that are radically unstable and inscrutable to ordinary speakers are not well suited to either the explanatory or the normative roles of semantic contents. The worry is that they are simply too fine-grained to be helpful in either (i) explaining how expressions have been used, or (ii) setting appropriate normative standards for how they should be used.

Stability of reference has been a central motivation in externalist explanations of epistemic progress both in science and in rational inquiry into the nature of commonsense kinds (sofa, spice, marriage, free will). The presupposition of this type of inquiry is that one is getting closer to the truth about the very same objects, kinds, or properties we were originally talking about. And in conversation with others, the presumption is that the interlocutors are talking about the reference, despite differences in their fine-grained understanding and use. An interpretation function that made the reference highly unstable through rational inquiry and debate would fail to explain these cases of apparent semantic engagement. And they would set normative standards for correctness of

---

[7]An additional ad hominem point. Contrary to what Cappelen suggests (67), we believe the arguments Dorr and Hawthorne (2014) offer for radical instability of reference ("plasticity") hinge on vagueness. However, we won't press this minor issue here.

token claims that would strike participants in rational inquiry and debate as wrongheaded and irrelevant. So there is good reason to think that a moderate externalist should accept a theory of interpretation that generally vindicates speakers' presuppositions of semantic coordination.

Second, an existing model: There is in fact a large philosophical literature in which externalists have addressed the question of how to reconcile undetected shifts in reference with privileged access to sameness and difference of thought contents. One way to sum up externalist solutions is this: an adequate interpretation of thought content must assign contents in such a way that a thinker's presumptions of coreference are reliable across a wide range of cases (e.g., cotemporal thoughts, occurrent memories, trains of reasoning, homophonic interpretations).[8] Theoretical trade-offs have to be made: in order to vindicate the presumption of semantic stability of memories over time, for instance, one may have to deny the presumption of semantic stability among certain cotemporal thoughts. But the key point is that an adequate theory of reference should find the *best overall balance* among these presumptions of coreference, consonant with other facts about understanding and use. The fact that not all of a thinker's presumptions of semantic stability can be vindicated does not mean that such presumptions should be ignored altogether. Vindicating the stability of reference over time and between individuals is one among a number of competing interpretive constraints in assigning semantic contents to an individual's thoughts. But given the importance of stability in vindicating rational inquiry and debate, there is good reason to suppose that these interpretative constraints will significantly limit the range of errors about stability of reference.

Moderate externalists about linguistic content can make a similar move. They can argue that any adequate semantic interpretation of particular uses of a term must make speakers' presumptions of coreference over time and between speakers reliable. Again, this interpretive constraint would be subject to trade-offs within an overall semantic theory. With such an interpretive constraint in place, tiny shifts in usage are unlikely to generate tiny inscrutable shifts in reference. Shared and relatively stable semantic contents would be better suited, we submit, to fulfilling the explanatory and normative roles of reference for linguistic content than highly unstable and inscrutable referential assignments advocated by Dorr and Hawthorne (2014). If we want the referential assignments to help in the rationalizing explanation of joint inquiry and debate or to set reasonable normative standards for the upshot of those processes, then stability of reference should be a constraint on interpretation. This constraint would not make presumptions of linguistic coreference infallible, but it could ground both the reliability and epistemic justification of the presumption of stability.

In sum, we agree that Cappelen's argument for (a), Unknown Boundaries, would be sound *if* he could give convincing reasons to suppose that reference must have sharp boundaries *or* that reference is highly sensitive to tiny shifts in usage. But neither position is entailed by moderate externalism, and both strike us as very implausible. In addition, we'd like to stress once again that conclusion (a) is too weak to support Cappelen's broader views about conceptual engineering. Even if sharp boundaries and tiny changes in extension are inscrutable, we may still know enough about the reference of our terms to evaluate and plan for long-term changes in reference. So even if he can establish (a), Cappelen wouldn't have shown that conceptual engineers are necessarily destined to grope in the dark.

In this section, we have considered two versions of Inscrutability, Cluelessness and Unknown Boundaries, which we took to be suggested by Cappelen's discussion of externalism. We have argued that moderate externalists should reject both. Moreover, if these theses are rejected, there seems to be little threat to knowledgeable conceptual engineering. If Cluelessness is false, then there's no reason to deny speakers have reliable knowledge of the reference. Similarly, if Unknown

---

[8]Theorists differ over the precise character of the interpretative constraint: whether it applies to all of a person's thoughts at a time, or to thoughts preserved in memory, or to thoughts that occur together in a stream of reasoning. And theorists also differ over whether this constraint is defeasible. See Schroeter (2007) for an overview of the literature and an argument that the presumption of coreference is fallible.

Boundaries is false, then there's no reason to deny that we can have reliable knowledge of coreference and change of reference. We conclude that much stronger arguments are needed to make the case that externalists are committed to a type of opacity of reference and reference change that jeopardizes the prospects of successful conceptual engineering.

2. Should we care about inscrutable reference?

Cappelen is aware of the controversial nature of his views about the inscrutability of reference. So he considers the following important objection:

> One might object that if extensions (and intensions) are inscrutable, incomprehensible, out of our control, etc., why think they play important roles in our cognitive lives? Surely such "hidden" meanings are irrelevant to anything we should care about. They are some kind of exotic creatures that have no genuine significance for us. (82–83).

The challenge, then, is to explain why we should care about the reference of our terms if that reference is forever cognitively inaccessible.

In response, Cappelen argues (i) that the precise extensions of our terms are not "hidden," and (ii) that we do, in fact, care about precise extensions. Here's the core of his response:

> We all know, for example that the extension of "belief" consists of all and only beliefs and that the extension of "person" consists of all and only persons. That kind of disquotational knowledge is true, informative, and accessible to us despite Inscrutability. So Inscrutability doesn't imply that we can't know what we're talking about. Accordingly, contrary to the objection, we are *not* ignorant of what, e.g., "person" denotes. We know *exactly* what's in its extension. And that matters to someone who wants to understand what persons are. If you are interested in what persons are, then you are interested in what is in the extension of "persons." (83)

This line of argument needs some unpacking.

Cappelen's first move is to invoke disquotation as a way of expressing one's knowledge of the precise extension of one's own terms. The key idea here is that any minimally competent speaker can *use* any term in her repertoire to pick out the precise extension of that term. Now, externalists have argued that this semantic competence is consistent with significant ignorance and error in your recognitional capacities or in your ability to provide a substantive definition or explication of the conditions for falling into the extension. Even with this sort of ignorance and error, you can still use your term to pick out the reference. For instance, even if you haven't studied physics, you might still count as linguistically competent with the term "gluon" because of you've heard authoritative physicists use it. In that case, you can knowledgeably assert:

> The extension of "gluon" consists of all and only gluons.

You may not be able explain what gluons are. All you really know about gluons is they are whatever those physicists talk about using that term (and even this claim might turn out to be not quite right). Still, according to Cappelen, you know *exactly* what gluons are because your use of the term "gluon," in fact, picks out all and only gluons. So your homophonic interpretation of the term's extension is guaranteed to be true whenever you utter it. And given your minimal semantic competence with the term and your knowledge of the disquotational schema, you can use this sentence to *knowledgably* attribute exactly the correct extension to your term. That's the sense in which you know the precise extension of each and every one of the terms in your repertoire. This disquotational knowledge, however, is perfectly consistent with "massive, fundamental" error about how to specify the reference independently of simply reusing the term in question.

So Cappelen's first point is that disquotational knowledge is cheap; we are guaranteed to have it even if a substantive specification of the extension of our term is inscrutable. This brings us to the second point. Why should we *care* about an inscrutable reference? Here, Cappelen simply appeals to the idea that *by hypothesis* your use of the term "gluon" picks out a particular inscrutable extension: when you use "gluon," you're talking about gluons—whatever those things may be. So whenever you formulate a question like "What are gluons after all?" you're raising a question about what it takes to fall into that particular inscrutable extension.

In sum, Cappelen's response to the mooted objection seems to take the following form:

1. Reference is inscrutable.
2. We care about the reference of our terms.
3. Therefore, we care about inscrutable reference.

Cappelen starts from the assumption that the inscrutability of reference is a corollary of moderate externalism. He then seeks to refute the objection that we don't care about inscrutable reference by pointing out that we do, in fact, care about the reference of our terms because reference fixes the subject matter of our "what is *x*?" questions. So we care about inscrutable reference because reference determines the goal for rational inquiry when we seek to answer "what is *x*?" questions.

However, this response begs the question at issue. The objection to Cappelen's position can be formulated roughly as follows:

a. We care about the reference of our terms.
b. We don't care about the outputs of an inscrutable interpretation function.
c. Therefore, the reference of our terms isn't determined by an inscrutable interpretation function.

Clearly, an adequate response to this objection cannot start from the premise that the conclusion c is false. In order to engage with the objection, Cappelen needs to address b, the key premise: why should we care about something inscrutable? In particular, why should rational inquiry be beholden to standards of correctness that are inscrutable?

The worry is that inscrutable referential assignments are not worth caring about because we could never be in a position to specify them in any nontrivial manner. We agree that when we engage in scientific theorizing about the nature of gluons, or philosophical theorizing about the nature of persons, our aim is to refine our substantive characterizations of the reference in such a way as to better approximate the truth. But consider the norms that ought to govern rational inquiry into these questions: we should not try to identify the kind of highly unstable intensions that Cappelen suggests. These shifting intensions are irrelevant to the broader practical and theoretical interests at stake when we inquire into the nature persons, e.g., the moral and legal status we accord to persons, the assumption that continuity of persons normally entails continuity of psychological traits, our metaphysical assumptions about what explains personhood, and so on. These are the kinds of consideration we take into account in rational inquiry into the nature of persons—not facts about what was said in the last five minutes. And even if, *per impossibile*, we were to hit on a specification of an inscrutable shifty extension, this would not count as success in rational inquiry as judged by our best reflective methods. Indeed, Cappelen suggests that our best methods of rational inquiry are just as likely to lead us *away* from a true specification of an inscrutable reference as toward it (159). So the challenge for Cappelen is to explain why we should care about unknowable shifty extensions he favors. Why suppose they determine the correct answers to the question "what is *x*"?

This way of formulating the objection appeals to our first-person perspective as rational inquirers. But a more systematic form of the objection will step back and consider whether inscrutable extensions can fulfill the key explanatory and normative roles of semantic contents. As we noted in the previous section, inscrutable semantic assignments seem ill suited to these two

types of role. If reference is inscrutable, an accurate answer to "what is *x*?" questions is forever beyond the cognitive reach of both competent speakers and theorists alike. If this is right, it's not at all clear why inscrutable extensions would set appropriate normative standards for evaluating the rationality or success of inquiry of ordinary speakers—even from the perspective of a theorist. Nor is it clear why these inscrutable extensions would be useful in theoretical explanations of the linguistic behavior of speakers and hearers: insofar as they are inscrutable, these referential assignments seem irrelevant to causally explaining individual or group behavior. So why should we, as semantic theorists, care about these inscrutable extensions? What theoretical role to the play for us?

Without a convincing explanation of why we *should* care about inscrutable extensions—either from the theoretical or participant perspectives—the objection against Cappelen stands.

3. Topic continuity

So far, we have focused on inscrutable reference. But Cappelen introduces a new theoretical notion called *continuity of topic*, which is not inscrutable in this way. The notion of continuity of topic is intended to capture our intuitions that we can use a term with the same meaning despite possible shifts in its extension. Sometimes, at least, to change the reference of a term is *not* to change the topic under discussion. Why isn't this always the case?

To illustrate the problem, Cappelen (95) cites Strawson's objection to Carnap's proposal to replace inexact concepts with exact ones. According to Strawson, this program rests on a mistake:

> [T]o offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of nonscientific discourse, is to do something utterly irrelevant—is a sheer misunderstanding, like offering a textbook on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart. […] typical philosophical problems about the concepts used in nonscientific discourse cannot be solved by laying down the rules of exact and fruitful concepts in science. To do this last is not to solve the typical philosophical problem, but to change the subject. (Strawson 1963, 503)

More generally, the worry is that if conceptual engineering changes the extension or intension of a term, it will *eo ipso* change the topic under discussion. Proponents of conceptual engineering are thus faced with the following challenge:

> How is conceptual revision (or amelioration or evolution) compatible with continuity of inquiry, continuity of interpretation, diachronic agreement and disagreement, and how does the activity avoid generating massive amounts of verbal disputes? (8)

Cappelen's response is that *topic continuity* is more "coarse-grained" than sameness of extension or intension. So two uses of "woman" that pick out different extensions may nonetheless address the same topic. [9]

Cappelen's central strategy for developing his account of topic continuity is to appeal to data about samesaying, i.e., "data about when we correctly describe people as *having said the same thing*" (107):

> […] A and B can samesay each other using a sentence "Fa," even though the extension of "F" in A's speech differs from the extension of "F" in B's speech. I'll illustrate this point in two ways: (i) by data about uses of context-sensitive expressions (and what I call intercontextual samesaying), and (ii) by thinking about diachronic samesaying. (107)

---

[9]According to Cappelen, conceptual engineering does not always involve continuity of topic. In those cases, Cappelen explains the point of conceptual engineering in terms of "lexical effects," in which we seek to change the extension but preserve brute psychological associations with a lexical item (122–34). We'll set aside this additional aspect of Cappelen's proposal here.

Samesaying is thus a matter of our willingness to use the same term in indirect reports of other's speech acts. Take, for instance, a context-sensitive expression like "tall." Even though the term's extension varies depending on the context in which it is uttered, we often disregard such variations in extension in indirect speech reports. We are happy to claim "A and B said Serena is tall" even though A's and B's utterances occurred in contexts which generated different extensions. Diachronic cases are similar in this respect: we're happy to use the term "marriage" to characterize what people said in the sixteenth century, when the term arguably had a very different extension. Cappelen's proposal, then, is to explain sameness of topic in terms of samesaying: "Sameness of topic goes hand in hand with samesaying (108)."

What exactly is this "hand in hand" relationship? Cappelen's suggestion seems to be that sameness of topic is not determined by our *actual* reporting practices, but rather by the reporting practices we *should* have (119). So the key question is a normative one: What makes an indirect speech report *correct*?

For normative factualists like ourselves, there can be genuine facts of the matter about which speech reports are correct and which are incorrect. In principle, a normative theory of correct speech reports seeks to specify how empirical facts determine whether a speech report is correct or incorrect. To address the issue of samesaying, a normative theory of *homophonic indirect speech reports* would need to say just how much the extension of the target speaker's use of a target term can vary from its extension in the reporter's own mouth and still yield a true report.

However, Cappelen does not offer a positive theory of the empirical conditions under which a speech report is correct.

> Insofar as I have a theory, it's best called *the Contestation theory of the Limits of Revision*: just as there are no fixed rules for how conceptual change can be implemented, there are no fixed rules for how far revision can go. The limits of revision are themselves up for revision, contestation, and negotiation. If there are any rules here at all, it's that we make up the rules along the way. (116)

What's crucial to understanding the correctness of speech reports is just that the standards that one speaker takes to govern speech reports can always be challenged by others. Cappelen takes claims about how we should report attitudes as "recommendations" (119). But these recommendations are not answerable to any objective adjudication; instead, we simply try to influence others via "contestation." Cappelen's position thus seems to be a form of *nonfactualism*: there are no objective facts—facts stable across different perspectives engaged in the process of contestation—about which indirect speech reports are correct. "Samesaying" is in the eye of the beholder. If this is right, then Cappelen seems to be proposing a nonfactualist explication of *continuity of topic*.

Be that as it may, we have a worry about Cappelen's general strategy of explaining *continuity of topic* in terms of the *norms governing indirect quotation*. The problem, in a nutshell, is that these two notions are answerable to different constraints. The correctness of indirect speech reports is—and should be—highly sensitive to facts about the reporter's own conversational context. But the judgments about continuity of topic should not be sensitive to the exigencies of the judger's own conversational context.

Topic continuity is supposed to set limits for rational inquiry and debate. For instance, sameness of topic determines whether two theorists engaging in a long-running debate about nature of "genes," or "marriage," or "free will" are talking about the same topic, or whether they are talking past each other. And it determines whether the answer one accepts at the end of rational inquiry is a response to the question one originally asked. On the face of it, then, sameness of topic depends entirely on facts about the people actually using the terms "gene," or "marriage," or "free will" and the empirical circumstances in which their discussion takes place. Facts about the interests or circumstances of external observers are simply irrelevant to setting normative standards for those engaged in rational inquiry and debate. Even from a third-person perspective, if you're trying to

determine whether some debate has coordinated on the same topic or whether the participants are talking past each other, you should be focusing on discerning the facts about the context of the discussion. The *rationality and successful epistemic coordination* of those engaged in a debate does not vary depending on whether we ourselves are talking to a child or an expert. In contrast, whether it's appropriate to use a homophonic sentence to characterize what was said is highly dependent on facts about your current conversational context: who you're speaking to, what your shared background assumptions are, which things are mutually salient in your current environment, etc. Whether an indirect speech report is correct thus depends essentially on facts that are entirely *extrinsic* to the original context in which the reported speech act occurred. But whether the interlocutors in that context were rationally engaged with each other depends on facts *intrinsic* to their context—facts about our own interlocutors and conversational interests are irrelevant.

Cappelen's explanation of sameness of topic seems to us to miss the point. Norms governing indirect speech reports are too ill disciplined and too sensitive to extraneous factors to capture the norms of rationality governing inquiry and debate.

4. Conclusion

In this paper, we have challenged Cappelen's claim that semantic facts are inscrutable. We distinguished two versions of the Inscrutability thesis, Unknown Boundaries and Cluelessness, and we argued that, *pace* Cappelen, moderate externalism is not committed to either one. Indeed, we've argued that most metasemantic theorists rightly reject these theses. So Cappelen has provided no compelling reason to think that we cannot know how changing patterns of use can change the reference of our terms—particularly in the case of long-term shifts in reference within a community. Thus, there is no reason to think that conceptual engineering is a matter of flailing in the dark. We've also argued that Cappelen's positive account of topic continuity fails to capture the norms governing inquiry and debate.

In closing, we'd like to highlight one final worry about Cappelen's Inscrutability thesis, which arises from his apparent embrace of semantic nonfactualism. As we noted, Cappelen's Contestation theory holds that there are *no fixed rules* setting limits on continuity of topic because all proposed limits are up for contestation. But this absence of fixed rules is not just confined to questions topic continuity. According to Cappelen, there are no fixed rules at the level of semantics and metasemantics either because semantic and metasemantic rules are "in flux" (69).

Cappelen starts by endorsing Williamson's (1994) claim that there is no "algorithm" that fixes semantic facts on the basis of empirical facts:

> I think this is exactly right and it is relevant in this context: if we're looking for an algorithm for how to change meanings, we are in effect asking for a recipe for extracting meaning from use and we have no good reason to think such a recipe exists. (67)

Although this may seem initially like an epistemic claim, he clarifies that he intends it as a metaphysical claim:

> Again, I suspect no algorithm exists—it's just that we haven't been able to articulate one, but there literally isn't one. (67)

The claim is there is no objective interpretation function. It follows that there are no objective facts about extensions or intensions.

To bring home the radical nature of his position, Cappelen (69, 154) approvingly cites this passage from Rorty:

> … philosophy is the greatest game of all precisely because it is the game of "changing the rules." This game can be won by attending to the patterns by which these rules are changed,

and formulating rules in terms of which to judge changes of rules. Those who take this view hold that philosophy in the old style—philosophy as "metaphysics, epistemology, and axiology"—needs to be replaced by metaphilosophy. Members of this school are, as it were, the metaphilosophers's metaphilosophers: since any metaphysical, epistemological or axiological arguments can be defeated by redefinition, nothing remains but to make a virtue of necessity and to study this process of redefinition itself. (Rorty 1961, 9)

On Cappelen's reading of Rorty, there are no fixed "rules of the game" in semantics or metasemantics. There is "a significant element of contestation" when we engage in debates about controversial cases like "marriage"; both the rules governing which things fall into the extensions of our terms *and* the rules governing why they do so are up for contestation and change (69). On the face of it, Cappelen seems to embrace a version of *nonfactualism* about semantics, metasemantics, and topic continuity. When it comes to meaning, it's contestation all the way down.

Our final worry for Cappelen is whether this nonfactualism is consistent with the inscrutability of reference. If there are no objective facts about the reference of your terms, then there is nothing *there* for you to have (or to fail to have) epistemic access to. Although it might seem as if the radical claim that there is no interpretation function would make the problem of epistemic access to reference even more severe, in fact the problem of epistemic access disappears. Reference isn't inscrutable for the simple reason that there are no facts about reference in the first place. The moral of the book, then, is *not* that conceptual engineers are groping in the dark hoping to chance upon an improved reference for our terms. If there are no objective facts about semantics and metasemantics, then *all* debates about meaning are conceptual engineering. The moral is that meaning is contestation all the way down.

Laura Schroeter, senior lecturer at the University of Melbourne, works primarily on concepts, reference determination, and metaethics.

François Schroeter is a senior lecturer at the University of Melbourne. His work focuses mainly on metaethics and moral psychology

References

Boyd, Richard. 1980. "Scientific Realism and Naturalistic Epistemology." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* Vol. 2: 613–62.

Boyd, Richard. 2013. "Semantic Externalism and Knowing Our Own Minds: Ignoring Twin-Earth and Doing Naturalistic Philosophy." *Theoria* 79: 204–28.

Burge, Tyler. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4: 73–121.

Burge, Tyler. 1986. "Intellectual Norms and Foundations of Mind." *Journal of Philosophy* 83: 697–720.

Burge, Tyler. 1988. "Individualism and Self-Knowledge." *Journal of Philosophy* 85: 649–63.

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

Chalmers, David. 2006. "The Foundations of Two-Dimensional Semantics." In *Two-Dimensional Semantics: Foundations and Applications*, edited by M. Garcia-Carpintero and J. Macia, 55–140. Oxford: Oxford University Press.

Dorr, Cian, and John Hawthorn. 2014. "Semantic Plasticity and Speech Reports." *Philosophical Review* 123: 281–338.

Evans, Gareth. 1973. "The Causal Theory of Names." *Proceedings of the Aristotelian Society*, Supp. 47: 187–208.

Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Lewis, David. 1974. "Radical Interpretation." *Synthèse* 23: 331–44.

Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Millikan, Ruth Garrett. 1993. "White Queen Psychology: or, The Last Myth of the Given." In *White Queen Psychology and Other Essays for Alice*, 279–363. Cambridge, MA: MIT Press.

Millikan, Ruth Garrett. 2017. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford: Oxford University Press.

Peacocke, Christopher. 1992. *A Study of Concepts*. Cambridge, MA: MIT Press.

Putnam, Hilary. 1975. "The Meaning of 'Meaning'." *Minnesota Studies in the Philosophy of Science* 7: 131–93.

Rorty, Richard. 1961. "Recent Metaphilosophy." *Review of Metaphysics* 15: 299–318.

Schroeter, Laura. 2007. The Illusion of Transparency. *Australasian Journal of Philosophy* 85: 597–618.

Schroeter, Laura, and François Schroeter. 2015. "Rationalizing Self-Interpretation." In *The Palgrave Handbook of Philosophical Methods*, edited by C. Daly, 419–47. London: Palgrave Macmillan.

Strawson, P. F. 1963. "Carnap's Views on Conceptual Systems versus Natural Languages in Analytic Philosophy." In *The Philosophy of Rudolph Carnap*, edited by P. A. Schilpp, 503–18. Chicago: Open Court.

Williamson, Timothy. 1994. *Vagueness.* New York: Routledge.

Williamson, Timothy. 2007. *The Philosophy of Philosophy.* Oxford: Blackwell.

Yablo, Stephen. 2008. "No Fool's Cold: Notes on Illusions of Possibility." In *Thoughts: Papers on Mind, Meaning, and Modality*, 151–70. Oxford: Oxford University Press.