Abstract

Truth by convention, once thought to be the foundation of a uniquely promising approach to explaining our access to the truth in nonempirical domains, is nowadays widely considered an absurdity. Its fall from grace has been due largely to the influence of an argument that can be sketched as follows: our linguistic conventions have the power to make it the case that a sentence expresses a particular proposition, but they can't by themselves generate truth; whether a given proposition is true—and so whether the sentence that expresses it is true—is a matter of what the world is like, which means it isn't a matter of convention alone. The consensus is that this argument is decisive against truth by convention. Strikingly, though, it has rarely been formulated with much precision. Here I provide a new rendering of the argument, one that reveals its structure and makes transparent just what assumptions it requires, and then I assess conventionalists' prospects for resisting each of those assumptions. I conclude that the consensus is mistaken: contrary to what is almost universally thought, there remains a promising way forward for the conventionalist project. Along the way, I clarify conventionalists' commitments by thinking about what truth by convention would need to be like in order for conventionalism to do the epistemological work it's intended to do.

1 Conventionalism and its discontents

Conventionalism about sentences of a given sort is the doctrine that those sentences, in some sense to be specified, are made true and false by, or owe their truth values wholly to, linguistic convention. It has traditionally been invoked by empiricists to explain our apparent ability, in certain domains, to form true beliefs with near-perfect reliability,

† This is an author-produced version of an article accepted for publication in Philosophical Studies following peer review. The version of record is available at link.springer.com/article/10.1007/s11098-018-1088-5.
despite not having observational access to those domains. And it’s easy to see why the doctrine is appealing: how we could possibly be in contact with (say) an abstract realm of mathematical objects is, for those of us with empiricist or naturalist sympathies, hopelessly mysterious, and by endorsing conventionalism about mathematics we eliminate the need to posit any such contact. After all, if truths in a given domain are true by linguistic convention, our access to those truths need not be explained in terms of contact with any part of the world, abstract or otherwise—it’s explicable just in terms of our linguistic competence.

Such, anyway, is the promise of conventionalism. But the story certainly needs some filling in. If conventionalism in a given domain is to be tenable as a solution to our epistemological mystery, a plausible account is needed of how (and in what sense) we can, just by adopting certain linguistic conventions, make it the case that claims in that domain have the particular truth values they do. Unless such an account is forthcoming, to endorse conventionalism is just to exchange one mystery for another.

In fact, the situation is worse than this. Certain objections have been taken to show that no satisfactory account of truth by convention is even possible, and as a result, conventionalism has largely been abandoned by mainstream analytic philosophy. Most famously, Quine, in his influential series of attacks on conventionalism (developed most fully in “Truth by Convention” (1936) and “Carnap and Logical Truth” (1960)), has argued both that explicit convention can’t be the source of all truth in any given domain, on pain of regress, and that the notion of implicit convention can’t really be made sense of—there’s no naturalistically respectable way to draw a distinction between those sentences we accept as a matter of implicit convention and those we merely take to be obvious. And he seems to be right about explicit convention. So if he’s right

1 Conventionalism has also been invoked to explain our knowledge of those domains, our ability to have justified beliefs about those domains, etc.—indeed, this is how it’s usually presented, as a way of developing an analytic theory of a priori knowledge or justification. But I’ve chosen to focus here on reliability, for the following reason: It may be possible to defend a thin conception according to which (e.g.) justified belief is easier to come by than we might have thought—see, e.g., the pragmatic view defended by Boghossian (2003a, 2003b), according to which what explains the justification of certain of our logical beliefs is (roughly) that we must have them in order to reason responsibly at all. If such a conception is available, conventionalism may not be needed to account for our ability to have justified beliefs in the relevant domains, since we may be able to provide an account without explaining how we manage to get at the truth in those domains. But it’s not clear that we can account for our reliability in any similar way, since no analogous thin conception of reliability seems to be available—to explain our reliability just is to explain how we manage to get at the truth. Focusing on reliability, then, is one way of bringing out what’s especially attractive about conventionalism as compared to its alternatives. (In fact, conventionalism is the only view that has ever provided any real hope of a satisfying naturalist-friendly explanation of our reliability in nonempirical domains, though a full defense of that claim is beyond the scope of my discussion here.)

2 The mystery is especially obvious in the case of mathematics, but even in nonabstract domains, we appear able to reliably form true beliefs about unobserved parts of the world. For example, we haven’t seen every vixen in the world, but we still believe, correctly, that all of them are foxes—and that every vixen that will ever exist will be a fox, and even that that it’s impossible for any vixen to fail to be a fox.

3 This isn’t to say that there are no contemporary conventionalists—Einheuser (2006), Glock (2003, 2008), Sidelle (1989, 2009, 2010), and Warren (2015a, 2015b, 2017), to name a few, have in recent years issued defenses of conventionalism in various domains. But it’s a distinctly minority position. (García-Carpintero and Pérez Otero (2009) also defend what they say is a version of conventionalism, but their view doesn’t count as conventionalist in the sense I’ve described here—they deny that any sentence can be true by convention alone. As a result, their view doesn’t solve the epistemological mystery conventionalism is intended to solve.)

4 For a thorough review of Quine’s argument here, see Warren’s (2017) recent discussion of the Quinean
about implicit convention as well, then conventionalism is indeed hopeless.

Notice, though, that his worry about implicit convention is a worry not just for truth by implicit convention but for any metasemantic theory according to which claims are sometimes accepted as a matter of implicit convention. And even opponents of conventionalism are often happy to endorse the claim that our willingness to accept certain claims as a matter of implicit convention plays some role in explaining how our words come to mean what they do. So, insofar as there's a problem here, it's not unique to conventionalism—many opponents of that doctrine are just as committed as conventionalists are to the existence of the distinction Quine is calling into question. This suggests that Quine's critique, for all its renown, can't be what accounts for conventionalism's poor reputation.

A look at the literature confirms this suspicion—the problem most often taken to be fatal for conventionalism isn't either of the ones pointed out by Quine. Conventionalism fails, according to its opponents, not for subtle Quinean reasons but simply because sentences, since they say things about the world, can never owe their truth values wholly to convention, implicit or otherwise. (Sentences about our linguistic conventions are, of course, exceptions. For instance, the truth value of the sentence “It's a convention among English speakers to be willing to apply the predicate ‘vixen’ to any object to which one is willing to apply both 'female' and 'fox'” surely is, in some sense, a matter of what conventions are in place. But we can set such sentences aside—interesting versions of conventionalism pertain not to claims about how language is used but to object-level sentences such as “All vixens are foxes”, “Unicorns either exist or do not exist”, or “1 + 1 = 2”.) Any sentence $S$, so the argument goes, owes its truth value in part to linguistic convention and in part to what the world is like: what it is for $S$ to be true is just for $S$ to express some true proposition $p$, and while it is a matter of convention that $S$ expresses $p$—what linguistic conventions can do is fix the meanings of our expressions—it's not a matter of convention whether $p$ is itself true. After all, $p$ is a proposition about the world, which means the truth value of $p$ is going to depend on what the world is like. The truth value of $S$, then, is also going to depend on what the world is like. So no sentence can be true by convention alone—“The world”, as Sider (2011: 101) puts it, "must also cooperate".

Take, for instance, the sentence "All vixens are foxes." It's entirely obvious that this sentence says something about the world. In particular, it says something about vixens—namely, that they're foxes. So if this sentence is true by convention, then what it says must be true by convention, which means that our linguistic conventions somehow have the case against truth by convention.

Furthermore, if there's a solution to be found, it's going to be available to conventionalists just as much as to anyone else. Horwich (1998), for example, is no conventionalist, but he's committed to there being a distinction between those dispositions that are meaning-constituting and those that aren't. His account of this distinction is that the meaning-constituting dispositions for a given word are the ones that are explanatorily fundamental, in the sense that they can explain all other dispositions to use that word. If something like this is right, then Quine's critique presents no problem, for conventionalists or anyone else.

Some theorists state explicitly that Quine's critique doesn't get to the heart of what's wrong with conventionalism. Sider (2011: 100), for instance, says that "Quine's argument does not go far enough" because it doesn't "challenge the very idea of something's being 'true by convention'"—after all, Quine concedes that it is possible, by explicit stipulation, to make some sentences true. Sober (2000) and Benacerraf (1973) express similar sentiments.
power to make it the case that vixens are in fact foxes, to make the world one way rather than another. But that seems absurd: the conventions of English do make it the case that “All vixens are foxes” says that all vixens are foxes, but it seems clear that whether all vixens are foxes is in no way a matter of the conventions of English or any other language. It’s just a matter of what vixens are like. And if that’s right, then “All vixens are foxes” can’t owe its truth wholly to convention.

Call this the objection from worldly fact. It’s the standard argument against conventionalism, and it’s generally taken to be decisive. Strikingly, though, the objection is rarely laid out in very much detail. What usually happens instead is that the objection is briefly sketched in much the way I’ve just sketched it above, and this sketch is taken to be sufficient to show that conventionalism is hopeless.

What’s more, when detailed argumentation is offered, it often fails to target any genuine conventionalist commitment. Lewy, for example, thinks conventionalists would endorse the following (absurd) thesis: that, necessarily, if (e.g.) ‘vixen’ has the same meaning as ‘male fox’, then all vixens are male foxes. So he provides a series of careful arguments against this thesis. But those arguments seem irrelevant, for actual conventionalists don’t in fact endorse the thesis. Nor, as Wright (1985) points out, are they committed to it merely in virtue of their conventionalism, despite what’s sometimes thought. And if that’s right, then Lewy’s critique simply misses its intended target.

The blame for misunderstandings of this sort, though, doesn’t lie entirely with the opponents of conventionalism. After all, the claim that a sentence is “made true by convention” or “true in virtue of convention” or that it “owes its truth wholly to convention”, though suggestive, is hardly precise, and conventionalists’ attempts to clarify their doctrine are often unhelpful. For example, Lewy’s primary target, Wisdom, represents his view as one on which the object-level proposition that (e.g.) a thing is a vixen just in case it’s a female fox “makes the same factual claims” as the metalinguistic proposition that ‘vixen’ has the same meaning as ‘female fox’ (see his 1938: 462–463). But unless significant further clarification is offered, this is misleading at best: for one thing, on any remotely intuitive understanding of what it is to make a factual claim, the metalinguistic proposition makes a factual claim about the meanings of English words that the object-level proposition doesn’t make. As a result, it’s difficult to see just what Wisdom is trying to say here. Lewy’s confusion is understandable.

This is all to say that there’s reason to be dissatisfied with the state of the debate over truth by convention: conventionalists, by and large, have failed to precisely formulate their view, and their opponents have in turn dismissed that view by appeal to an objection that is itself not precisely formulated (and whose precisifications seem to be based on misunderstandings of what the view entails).

---

7 The objection (or something very much like it) has been endorsed in print by Sider (2011), Russell (2008), Williamson (2007), Hale (2002), Sober (2000), BonJour (1998), Horwich (1998), Boghossian (1996), Yablo (1992), Lewy (1976), Benacerraf (1973), David Lewis (1969), Pap (1958), Kneale (1947), C. I. Lewis (1946), and Ewing (1940), among others. Note, too, that neo-Fregeanism about arithmetic (see, e.g., Wright 1983), Schiffer’s (2003) theory of pleonastic entities, and other flavors of what Thomason (2015) calls “easy ontology” are often thought to be vulnerable to an objection of the same sort: that such approaches are workable only on the absurd view that we have the power to conjure objects into existence via our choice of language. Versions of this latter objection have been endorsed by van Inwagen (2016), Bennett (2009), Chalmers (2009), Boolos (1997), and Field (1984), among others.

8 I discuss this matter in some detail in §5 below.
My aim, then, is to do what I can to clarify the situation. In particular, what I'm going to do is produce as precise (and sympathetic) a reconstruction of the objection from worldly fact as I can, with a view to working out exactly what premises it depends on and whether conventionalists are in fact committed to those premises.

My contention, just so that it's clear where we're headed, is that things shake out as follows. Though the objection from worldly fact is not as straightforward as its proponents have taken it to be, this much is correct: conventionalists, despite what some of them have said, are committed to the thesis that, in a sense, linguistic conventions have the power to make the world one way rather than another. Whether (e.g.) all vixens are foxes does turn out to be a matter of convention. And this thesis is unorthodox, to be sure. But, properly understood, it isn't obviously absurd, for reasons I'll explain. And if that's right, then conventionalism remains worthy of investigation. This is what I'll be arguing here.

But first things first: in order to competently evaluate conventionalism's prospects in the face of the objection from worldly fact, we've got to have some grip on what conventionalists are and are not committed to. So I begin with an attempt to say something clarifying about how conventionalism is to be understood.

2 What conventionalism is and what it needs to be

As I suggested in my discussion of Wisdom and Lewy, the precise content of the conventionalist doctrine is elusive, even for conventionalists themselves. Ayer, for example, tries to elucidate his conventionalism about analytic sentences by claiming variously that they

- "make no statement whose truth can be accepted or denied" but "merely lay down a rule which can be followed or disobeyed" (1936b: 20);
- "are entirely devoid of factual content", since they don't "provide any information about any matter of fact" (1936a: 104);
- are "not…about 'things' at all, but simply about words" (1936a: 64).

These formulations, though, are all unfortunate, and not just because they're (at least seemingly) incompatible with one another. The third is similar to Wisdom's claim, and problematic for similar reasons: what the sentence "All vixens are foxes" says is just that all vixens are foxes, and this claim, on any remotely intuitive understanding of what it is for a claim to be about something, is about vixens, not words. And the first and second, if taken literally, respectively entail that it's not true that all vixens are foxes and that it's not a fact that all vixens are foxes. If conventionalists were really committed to these absurdities, conventionalism would be an easy doctrine to dismiss.

But conventionalists need not endorse any of these claims. They can say instead that, while what Ayer says is literally false—it's (of course) true that all vixens are foxes, and

---

9 He later decides that this view is untenable; see his 1946: 17.
10 See also, e.g., Malcolm 1940: 200 and Glock 2008: §2.
11 See also, e.g., Carnap 1934/1937: §74. Note, too, that Ayer later disavows any view on which this claim is taken literally, saying that, although he "sometimes seem[s] to imply that" analytic sentences "describe the way in which certain symbols are used", he neither "wish[es] to hold" that position nor "think[es] that [he is] committed to it" (1946: 16).
this is (of course) a fact about vixens—what he's trying to give voice to is just the idea that, in some sense to be specified, sentences like "All vixens are foxes" don't say anything substantive about what the world is like. All he's really committed to, then, is that, although the sentence "All vixens are foxes" may indeed be true, it doesn't say anything substantive—there's something degenerate about it.

Unsatisfactory formulations aside, this idea, vague though it is, does seem to be what Ayer and his fellow conventionalists are struggling to articulate. And it's an intuitively appealing idea, which may explain why, despite the fact that few contemporary theorists actually accept conventionalism, certain sorts of sentences (namely, logical and analytic truths) are very commonly characterized as "empty", "trivial", "true by definition", and the like. But it is, again, just a vague idea. The trick is turning it into a precise doctrine.

Full disclosure: I won't here be even attempting to pull off this trick. That said, I do want to try to shed some light on what it is for a true sentence to fail to say anything substantive, on what's supposed to be degenerate about (e.g.) the sentence "All vixens are foxes"—any clarity we can gain here is going to bring us closer to understanding just what's distinctive in the conventionalist position. So: what is it that's supposed to be lacking about "All vixens are foxes", as compared to (e.g.) the sentence "All vixens weigh less than a ton"?

As we search for an answer to this question, we can focus our investigation by keeping in mind the epistemological purpose of conventionalism. Conventionalists' primary goal, recall, is to dissolve a puzzle, to explain how we manage to reliably believe truths in domains our access to which would otherwise be, by naturalist lights, hopelessly mysterious. So, as we try to understand what conventionalism comes to, we can immediately reject as inadequate any interpretation on which the doctrine doesn't offer at least some reasonable hope of dissolving this puzzle. In other words, in order for conventionalism to serve its purpose, the following constraint must be met:

**Naturalist-Friendliness.** An adequate answer to our epistemological puzzle in a given domain must give us the resources to explain—without appeal to any facts our nonobservational access to which remains mysterious by naturalist lights—how we manage to reliably believe truths in that domain.

If we want to have any chance of being accurate to conventionalists' intentions, then, we must construe the thesis that certain true sentences say nothing substantive in such a way that it would make sense, given the Naturalist-Friendliness constraint, for conventionalists to make that thesis the basis of their view. Or, in other words: as we try to nail down what the conventionalist doctrine is, we should think about what that doctrine would need to be in order for conventionalism to serve its epistemological purpose.

What all this suggests is that a true sentence that says nothing substantive must, for conventionalists, be one whose truth can in some sense be fully explained by appeal only to facts our access to which is not mysterious by naturalist lights. And it's obvious, I take it, that the facts that are going to be relevant, according to conventionalists, are facts

---

12 For more on contemporary theorists' failure to purge their thinking of conventionalist metaphor, see Sider 2011: §6.5.
about our linguistic conventions. For conventionalists, then, a sentence that's true by convention must be one whose truth is fully explicable just by appeal to facts about our linguistic conventions. And if that's right, then perhaps conventionalism is, as Warren (2015b: 88) suggests, best understood as a doctrine about explanation: “What is distinctive of conventionalism about some branch of discourse D is that linguistic conventions are supposed to fully explain the truth of sentences in D”.

This is still not fully precise, of course—the notion of explanation is itself notoriously difficult to pin down. But we're making some progress. And furthermore, the language of explanation makes additional progress possible: even without taking a stand as to the correct theory of explanation, we can say with some confidence that there are certain sorts of explanation that conventionalists don't have in mind. First, the sort of explanation relevant to conventionalism must be nonpragmatic, in the sense that what counts as a good explanation isn't dependent on any thinker's background beliefs, interests, etc.—otherwise, conventionalism might turn out to be true relative to some background beliefs and false relative to others, in which case it couldn't play the foundational epistemological role required of it. And second, it's reasonably clear that the relevant sort of explanation is not causal explanation—the relationship between our linguistic conventions and the truth of certain sentences doesn't seem to have the right features to be causal. (For instance, effects are generally taken to come after their causes, but the conventionalist view certainly isn't that we adopt some conventions and then, at some later time, certain sentences come to be true.)

Conventionalism in a given domain, then, can be understood as the doctrine that our linguistic conventions fully explain, in some nonpragmatic and noncausal way, the truth of sentences in that domain. But given this restriction to nonpragmatic and noncausal explanation, we might be inclined to wonder just what sort of explanation conventionalists do have in mind. Most of the usual suspects have been ruled out, after all; what else is left?

As it turns out, there's a family of nonpragmatic, noncausal relations of explanatory dependence that have been the object of much philosophical attention in recent years: grounding relations. Facts about grounding are generally taken to be facts about what's explanatorily prior to what, in some metaphysically robust but noncausal sense—as Fine (2001: 15) puts it, “If the truth that P is grounded in other truths, then they account for its truth; P's being the case holds in virtue of the other truths' being the case”. And this sounds a lot like what conventionalists are after.

In fact, on certain views, to be a grounding relation just is to be the kind of relation that figures in noncausal explanations. Audi's (2012: 688) view, for instance, is like this: for him, all it is for something to be a grounding relation is for it to be a “noncausal relation of determination” whose relata are facts. (His argument for the existence of a grounding relation is essentially just that one fact can explain another without being a cause of it—i.e., that one fact can determine another in the way required for the former to explain the latter even if the two facts aren't causally related.) And if such a view is correct—if noncausal determination is all that's required for one fact to ground

---

13 There are, of course, questions to be answered here about just how it is that we have access to facts about our conventions. But there's no bar in principle to a naturalist-friendly way of answering those questions.

14 Though this, like most claims about causation, is controversial.
another—then conventionalism certainly is a doctrine about what grounds what.

That said, if conventionalism is a doctrine about what grounds what, then, for conventionalism to have any hope of doing the epistemological work it’s intended to do, there must be something deeply wrong with standard theories of the nature of grounding. On standard theories, after all, facts about what grounds what are facts about the metaphysical structure of reality, which means that, if conventionalism is a view about what the grounding facts are, then conventionalists’ explanation of how we can access the truth about whether (e.g.) all vixens are foxes requires appeal to facts about the structure of reality. But in that case, conventionalism can meet the Naturalist-Friendliness constraint only if there’s some naturalist-friendly explanation available of our nonobservational access to facts about the structure of reality. And since no such explanation is available, standard theories of grounding entail that conventionalism, interpreted as a doctrine about what grounds what, can’t do the epistemological work it’s intended to do.

Conventionalists, then, must do one of the following two things: deny that conventionalism is a view about grounding (and so insist that not all noncausal determination relations between facts are grounding relations) or provide some argument showing that standard views of the nature of grounding are incorrect. And we’d need to do quite a bit more work on the nature of grounding to determine which of these tacks conventionalists should take. So, for the purposes of this discussion, at least, I’ll leave this question to the side—I won’t describe the explanatory relations conventionalists are interested in as grounding relations, but I also won’t reject outright the claim that these relations are grounding relations.

In any case, though, these relations—whether they’re grounding relations or not—are going to be quite formally similar to grounding relations. After all, even if the two families of relations are nonidentical, both are families of noncausal, nonpragmatic relations of explanatory dependence. And in the recent literature on grounding, the formal structure of the various grounding relations has been explored in some detail (in, e.g., Rosen 2010, Audi 2012, and Fine 2012a, 2012b). So our investigation into conventionalism will inevitably contain echoes of this literature—again, the formal structure of the explanatory relations conventionalists are interested in is going to mirror the formal structure of the grounding relations, even if it turns out that these two kinds of relation are distinct.

Summing up: by thinking about what epistemological work conventionalism is intended to do, we’ve arrived at the conclusion that conventionalism (in a given domain) is to be interpreted as the doctrine that the truth of sentences (in that domain) is fully explained, in some nonpragmatic, noncausal way, by our linguistic conventions (where this explanatory relation may or may not be a grounding relation). This, I recognize, is unsatisfying—we haven’t said exactly what sort of explanation is in play here, and so we haven’t made the conventionalist doctrine fully precise. But as I said, I’m not even attempting here to make that doctrine fully precise. I’m just attempting to provide some clarity as to what’s distinctive about conventionalism. And at least this much, I think,
has been done. (I take it that the way to gain further clarity here is to generate a full metasemantic theory that can dissolve the epistemological puzzle motivating conventionalism and then to think about what sorts of explanatory claims are entailed by that theory. But this is well beyond the scope of the present discussion.)

It will be useful, before we move on to our discussion of the objection from worldly fact, to emphasize one other feature of the epistemological problem conventionalism is intended to solve. The point of conventionalism, once again, is to explain how we manage to reliably form true beliefs in certain domains to which we don't have observational access. But what needs to be explained is not, or at least is not only, the reliability of our metalinguistic beliefs about the truth values of sentences. It's not enough to explain our ability to reliably form true beliefs about whether sentences such as “All vixens are foxes” are true; after all, we might believe that the relevant sentences are true despite having no idea what they say. (I might, for instance, believe that the Spanish sentence “Ningún soltero es un hombre casado” is true without knowing what it says, just on the basis of the testimony of a trustworthy bilingual friend.) What we're really interested in is our ability to reliably form true beliefs about whether (e.g.) all vixens are foxes. That is, for conventionalism (or any rival view) to serve its purpose, it must meet the following requirement:

**Object-Level Belief.** An adequate answer to our epistemological puzzle in a given domain must give us the resources to explain our ability to reliably form true object-level beliefs—not just true metalinguistic beliefs—in that domain.

Otherwise, our solution to the epistemological problem motivating conventionalism will be incomplete at best.

Both Object-Level Belief and Naturalist-Friendliness will have roles to play in the below discussion of the objection from worldly fact. For now, though, I want to end this section with a brief demonstration of how these constraints can be used to show that conventionalism is well motivated in the first place.

Some theorists have suggested that there's no reason to endorse conventionalism: it's possible (they say) to provide a different sort of analytic theory of the a priori, one that doesn't rely on anything as radical as truth by convention. In particular, they suggest that, though no sentences are true by convention, we can fix the meanings of as-yet-undefined expressions by stipulating the truth of certain sentences containing those expressions (or by otherwise accepting those sentences as true as a matter of convention; explicit stipulation isn't what's important here). The story is roughly as follows: when we stipulate that certain sentences containing some expression \( E \) are to count as true (or when we accept them as true as a matter of convention), we thereby pick out among all the possible meanings for \( E \) (where these possible meanings exist independently of our stipulations) whatever one it needs to have for those sentences to be true.\(^\text{17}\) On a view of this sort, the sentences in question aren't true by convention alone—they turn out to be true only on the condition that the world is able to provide us with some appropriate meaning for \( E \). If the world doesn't cooperate at least to this extent, then we'll have failed

\(^{17}\text{Proponents of views of this sort include Boghossian (1996) and Sider (2011).}\)
to fix any meaning for $E$ at all, despite our best efforts. Nevertheless, there’s some temptation to insist here that our ability to reliably form true beliefs in the relevant domains has been explained, since—when we do successfully fix a meaning for $E$—we believe, correctly, that the relevant sentences are true. And if that’s right, then at least some of naturalists’ worries go away, despite the fact that views like this are far less radical than conventionalism is. Our question, then, is whether such views can completely dissolve the epistemological puzzle motivating conventionalism (and thereby render such a radical doctrine unnecessary).

It turns out that they can’t—no view like the one under discussion can meet both the Object-Level Belief constraint and the Naturalist-Friendliness constraint at the same time. After all, even when we correctly believe that we’ve successfully given $E$ a meaning, what this tells us immediately is just that the relevant sentences are true. Suppose, for instance, that we’ve fixed the meaning of ‘vixen’ by stipulating the truth of the sentence “All vixens are foxes”. Then, given that our stipulation is successful, we can be sure that this sentence is true—we’ve stipulated as much. But this isn’t enough. If the view here is to meet Object-Level Belief, it needs to give us the resources to explain our ability to get from this correct metalinguistic belief to the correct object-level belief that all vixens are foxes. And to do that, it needs to give us the resources to explain how we come to understand the sentence “All vixens are foxes”. This is where the problem arises: To understand a sentence is to know its meaning, and on the view under discussion, meanings are things that exist independently of our stipulations. So we need an explanation of our ability to reliably form true beliefs about the meanings of sentences whose truth we’ve stipulated—of how we come to correctly believe (e.g.) that, of all the preexisting meanings that are available, this one is the one that’s attached to the sentence “All vixens are foxes”. And the view itself doesn’t give us the resources to explain this ability—the view itself, in fact, doesn’t give us the resources to explain how we can reliably form true beliefs about what meanings are available, how many meanings there are, or even whether there are any meanings at all. And that means that, when we appeal to our understanding of “All vixens are foxes” to explain how we come to correctly believe that all vixens are foxes, we appeal to a fact our access to which we have no way of explaining. So the view under discussion, in order to meet Object-Level Belief, must violate Naturalist-Friendliness.

Our epistemological puzzle, then, does provide genuine motivation for conventionalism—views that are less radical, such as the sort of view we’ve been discussing, are unable to do the explanatory work that needs to be done. So it’s important, if we’re interested in the prospects for dissolving our puzzle, to determine whether conventionalism

---

18 For a detailed critical discussion of views of this sort in the light of this possibility, see Horwich 1997 (or the revised version that appears as Horwich 1998: chap. 6).

19 For what it’s worth, I take this claim to be incorrect. What needs explanation, again, is our near-perfect reliability in the relevant domains, and in order to explain that, we need to explain our ability to be nearly perfectly reliable about when a meaning has successfully been given to $E$. And it’s not clear, on a view like this, that we have any way of explaining the latter ability. But we can set this worry aside for the sake of argument.

20 Hale and Wright (2000: 294) call this “the understanding problem” and take it to be a conclusive objection to the sort of view under discussion here.

21 For an interesting exchange regarding issues of this sort, see Boghossian’s treatment of implicit definition in his 1996, Glüer’s objection to Boghossian in her 2003, Boghossian’s response to Glüer in his 2003b, and Jenkins’s reply to Boghossian in her 2008.
remains tenable in the face of the objection from worldly fact. And to do that, we need a better understanding of how that objection works.

3 Reconstructing the objection from worldly fact

Recall: opponents of conventionalism are willing to allow that our sentences owe their truth values partly to our linguistic conventions; what they deny is the (much stronger) conventionalist doctrine that there are (ordinary, object-level) sentences that owe their truth values wholly to our conventions. Or, in the language of explanation: while opponents of conventionalism grant that facts about what linguistic conventions are in place play some role in explaining the truth (and falsity) of our sentences, they insist that, nevertheless, there’s no sentence whose truth (or falsity) is fully explained by these facts about convention. This is the conclusion the objection from worldly fact is intended to deliver. In this section I reconstruct that objection.

We can start by stating the intended conclusion a bit more precisely, as follows: there are no facts about what linguistic conventions are in place such that the truth of \( S \) (where \( S \) is any ordinary, object-level sentence) is fully explained by those facts. Or again, where \( C \) is the set of all facts about convention:

\[
\forall \Gamma (\Gamma \subseteq C \rightarrow \neg (S \text{ is true purely in virtue of } \Gamma))
\]

(Terminological note: when some facts play a role in explaining another fact, I’ll say that the latter fact obtains (at least) partly in virtue of the former ones, and when some facts fully explain another fact, I’ll say that the latter fact obtains purely in virtue of the former ones.\(^{22}\))

This conclusion, of course, is intended to apply to any (ordinary, object-level) sentence whatsoever. But in reconstructing the argument, it will be helpful, for concreteness, to concentrate on a particular sentence. So let \( S \) be “All vixens are foxes”.

The argument for \((C)\), then, proceeds by appeal to the following claim: whether all vixens are foxes isn’t a matter of the conventions of English or any other language. That is, the fact that all vixens are foxes is not itself a fact about what linguistic conventions are in place, nor can it be fully explained by any such facts.\(^{23}\) Or, a bit more formally,

\(^{22}\)I want to note that these in-virtue-of relations, whether identical to grounding relations or not, are at the very least formally similar to grounding relations. Their logic, on my view, resembles (but isn’t identical to) the logic of the relations of strict partial ground and strict full ground, respectively, in Fine’s (2012a, 2012b) pure logic of ground. (For more details, see my manuscript, “Metaphysical Analyticity and the Logic of ‘In Virtue Of’.”)

\(^{23}\)Here I’m casting the objection in terms of the fact that all vixens are foxes, but it could just as easily be cast in terms of (the truth of) the proposition that all vixens are foxes. (Indeed, it usually is cast in terms of the proposition.) I’ve chosen to appeal to the fact rather than to the proposition simply because talk of facts, unlike talk of propositions, is common in nonphilosophical contexts and so can’t be dismissed as mere philosophical extravagance. As I suggested in §2, conventionalism would be an easy doctrine to dismiss if it required us to deny that it’s a fact that all vixens are foxes.

All that said, the objection, I take it, is in the end just as strong when cast in terms of the proposition as it is when cast in terms of the fact. After all, skepticism about the existence of abstracta such as propositions seems unwarranted from a conventionalist perspective—one of the points of conventionalism, as I suggested in §1, is to deflate claims about abstracta and so to give naturalists a way of vindicating our claims to knowledge in abstract domains such as mathematics. (Recall that Carnap’s (1950/1956) attempt to allay skeptical wor-
(1) all vixens' being foxes ∈ C
(2) ∀Γ(Γ ⊆ C → ¬(all vixens are foxes purely in virtue of Γ))

And it's clear enough that (1), at least, is true—after all, the fact that all vixens are foxes is a fact about how things are with vixens, not a fact about how things are with us. So it isn't itself a fact about what conventions are in place in our linguistic community.

The substantive premise here, then, is (2), the claim that the fact that all vixens are foxes doesn't obtain purely in virtue of facts about convention. Proponents of the objection from worldly fact tend to be explicit about their commitment to this claim. And it's easy to see why: prima facie, it's exceedingly plausible that facts about parts of the world outside ourselves (including the fact that all vixens are foxes) do not depend entirely on us. To think otherwise would, it seems, be to embrace idealism.

Now: neither (1) nor (2) tells us anything at all about S. So a third premise is needed, one that says what the relationship is between the truth of S and the fact that all vixens are foxes. And while proponents of the objection from worldly fact don't tend to be explicit here, it's fairly clear that they're committed to the standard view of this relationship, according to which the fact plays a role in explaining the truth of the sentence. That is, they're committed to the following:

(3) S is true partly in virtue of all vixens' being foxes

So we can include this as our third premise.

The intended argument, then, seems to go as follows. Let Γ be some set of facts that, taken together, fully explain the truth of S. That is, let Γ be any set of facts such that S is true purely in virtue of Γ. By (3), we know that the fact that all vixens are foxes—call it v—has some role to play in explaining the truth of S. So it's plausible that any full explanation of the truth of S will appeal either to v itself or to some other facts that can themselves explain v. That is, Γ must contain either v or some facts that, taken together, fully explain v. And in either case, Γ contains at least one fact not about convention: by (1), v itself isn't a fact about convention, and by (2), any set of facts that can fully explain v must contain at least one fact not about convention. So any full explanation of the truth of S will appeal to at least one fact not about convention, which is just to say that (C) is true.

Most of the inferences in the above argument rely only on basic logic, but there's one exception: the inference from (3)—i.e., the claim that S is true at least partly in virtue of v, the fact that all vixens are foxes—to the claim that any full explanation of the truth of S will appeal either to v itself or to some other facts that can fully explain v. So it's worth taking a moment to think about whether there's reason to accept this inference.

ries about abstracta includes discussion of propositions as well as of numbers and classes.) Conventionalists shouldn't be in the business of rejecting out of hand abstracta that may prove theoretically useful.

24Yablo (1992: 878), for instance, says that whether a given proposition is true is a question "to which the rules of usage are quite irrelevant"; and Sober (2000: 247) says, similarly, that "the proposition expressed by [an English-language analytic sentence] does not depend for its truth on how English works".

25But stay tuned.

26To take one example, the following rhetorical question, which Boghossian (1996: 364) asks in the course of sketching his version of the objection from worldly fact, presupposes that a sentence has its truth value at least partly in virtue of a corresponding fact: "How could the mere fact that S means that p make it the case that S is true? Doesn't it also have to be the case that p?"
It turns out that there is. Given certain plausible assumptions about the logic of the in-virtue-of relations—assumptions to which conventionalists and their opponents are both committed—the inference is a good one. Here I explain why that is.

The first assumption needed is this: a fact plays a role in explaining another fact just in case there's some full explanation in which it plays a role. That is, a fact \( f \) obtains partly in virtue of a fact \( g \) just in case, for some \( \Gamma \) such that \( f \) obtains purely in virtue of \( \Gamma \), \( g \in \Gamma \). And this assumption seems trivial—to have a role to play in explaining some fact is surely just to have a role to play in some full explanation of that fact.\(^{27}\)

The second assumption, though, is more complicated. To see what it is, we need to note, first of all, that the facts relevant to explaining a given fact are going to form an explanatory hierarchy. To oversimplify a bit: certain facts directly explain the target fact, certain other facts directly explain the target fact, and so on. So we can give different (full) explanations of the target fact by appealing to facts at different levels in this hierarchy.\(^{28}\) Suppose, for instance, that the sky is colored purely in virtue of its being blue and that it's blue purely in virtue of its being cerulean. Then it may also be that the sky is colored purely in virtue of its being cerulean—it may be that the fact that the sky is cerulean, by explaining the fact that it's blue, thereby also explains the fact that it's colored.\(^{29}\) Or, more generally: since we can move up and down the explanatory hierarchy, it's possible for there to be two distinct sets of facts \( \Gamma \) and \( \Delta \) such that some fact obtains purely in virtue of \( \Gamma \) and also obtains purely in virtue of \( \Delta \).

Our second assumption, then, can be stated roughly as follows: in cases where there are two distinct such sets of facts, there can be no differences between \( \Gamma \) and \( \Delta \) except those that result from moving up and down the explanatory hierarchy. That is, there's, in a sense, only one explanatory story to be told about why a given fact obtains; there's no explanatory overdetermination.

Conventionalists and their opponents both accept this assumption, at least tacitly. The influence of the objection from worldly fact makes this clear: after all, if explanatory overdetermination were possible, then it would be possible for a sentence to be true purely in virtue of convention and to also be true in virtue of facts that are entirely unrelated to convention. In this case, the objection from worldly fact would be based on an obvious mistake—the objection requires us, after all, to conclude, on the basis of the claim that a sentence is true in virtue of nonconventional facts, that the sentence is not true purely in virtue of convention. And this is easy to spot even on cursory examination. Since no one has brought up this point, it seems clear that all parties to the discussion are committed to the assumption that explanatory overdetermination isn't

\(^{27}\) Indeed, in developing his notions of full ground and partial ground, Fine (2012a: 50) defines partial ground in terms of full ground using just such a biconditional. I don't define the partly-in-virtue-of relation in terms of the purely-in-virtue-of relation, but the biconditional does turn out to hold in the logical system I've developed.

\(^{28}\) One point of clarification about the notion of full explanation that's in play here: for an explanation to count as full in the relevant sense, it's not required that the explanatory story be traced all the way back to fundamental facts (if there are such things). On conventionalism, remember, the truth of certain sentences is fully explained by facts about what conventions are in place. But conventionalists certainly don't endorse the (absurd) claim that facts about convention are fundamental—it's abundantly clear that those facts are to be explained by facts about the linguistic dispositions, brain states, etc., of members of the linguistic community.

\(^{29}\) In fact, given that the purely-in-virtue-of relation is transitive, it turns out to be certainly true that the fact that the sky is cerulean fully explains the fact that it's colored.
possible.\(^{30}\)

It still remains to show that the two assumptions I’ve just described can justify the inference in question. We begin, recall, with \((3)\), according to which the truth of \(\text{S}\) is to be explained in part by the fact that all vixens are foxes—i.e., by \(v\). By the first assumption, there are some facts \(\Delta\) such that \(v \in \Delta\) and \(\text{S}\) is true purely in virtue of \(\Delta\). And by the second assumption, \text{any} full explanation of the truth of \(\text{S}\) will differ from \(\Delta\) only in ways that result from moving up or down the explanatory hierarchy. So, since \(\Gamma\) fully explains the truth of \(\text{S}\), \(\Gamma\) must contain one of the following: \(v\) itself, some facts that fully explain \(v\), or some fact that \(v\) plays a role in explaining and that itself plays a role in explaining the truth of \(\text{S}\). In order to justify our inference, then, we need only rule out the third of these possibilities.

Note, further, that on the standard view of the relationship between \(v\) (i.e., the fact that all vixens are foxes) and the fact that \(\text{S}\) is true, there just aren’t any facts between these two in the explanatory hierarchy. Instead, the fact that all vixens are foxes and the fact that \(\text{S}\) says that all vixens are foxes, taken together, constitute a full, direct explanation of the truth of \(\text{S}\). And if this is right, the third possibility can be ruled out: if there doesn’t even exist any fact that \(v\) plays a role in explaining and that itself plays a role in explaining the truth of \(\text{S}\), then \(\Gamma\) certainly doesn’t contain any such fact. Our inference, then, turns out to be a good one if we add the following to our premise set:

\[
(4) \quad \neg \exists f (\text{S is true partly in virtue of } f \land f \text{ obtains partly in virtue of all vixens’ being foxes})
\]

And it’s fairly clear that proponents of the objection from worldly fact are committed to this premise—again, it just follows from the standard view of the relationship between the truth of \(\text{S}\) and the fact that all vixens are foxes, and proponents of the objection, as I’ve suggested, are committed to that view.

Furthermore, even independently of its following from the standard view of the relationship between the relevant facts, this premise seems unproblematic—I’m not aware of \text{any} view according to which all vixens’ being foxes plays a role in explaining some other fact that plays a role in explaining the truth of \(\text{S}\), and it’s hard to imagine what a plausible view of this kind might look like.

So we’ve got a reasonable first pass at a reconstruction of the objection from worldly fact: it gets us the intended conclusion, \((C)\), via an argument from \((1)\), \((2)\), \((3)\), and \((4)\), all of which plausibly do play a role in the objection as usually sketched, and furthermore, the argument is valid given our two assumptions about the logic of the in-virtue-of relations. So this is the version of the objection we’ll be working with, at least for the moment.

Now: as I said, the argument here is valid. And conventionalists, in order to hold on to their view, need to resist the conclusion, which means they need to find a tenable theory on which at least one of the premises fails to be true. In the next section I consider strategies for doing so.

\(^{30}\)This, I want to note, is one place where the logic of the in-virtue-of relations under discussion here differs from Fine’s pure logic of ground.
4 Avenues of resistance

Our first order of business here is to decide which of the argument’s four premises conventionalists ought to consider rejecting. And we can set one of the possibilities aside immediately. As I’ve suggested, (1) is clearly true—on any remotely intuitive understanding of what it is for a fact to be about something, the fact that all vixens are foxes is a fact about what vixens are like, not a fact about what conventions are in place in our linguistic community.

It’s also relatively easy to set aside the possibility of rejecting (4). As I said above, this premise seems unproblematic—it’s difficult to imagine a plausible view on which it’s false. And even if such a view were available, it would be no help to conventionalists, for the following reason: the argument turns out to remain valid even if we replace (4) with the strictly weaker claim that there’s no fact about convention that lies between the truth of S and the fact that all vixens are foxes in the explanatory hierarchy, and conventionalists themselves, as it turns out, definitively are committed to this weaker claim.31 Given all that, trying to avoid (C) by rejecting (4) is, for conventionalists, a nonstarter.

So conventionalists must find a way to reject either (2), which says that facts about convention don’t fully explain the fact that all vixens are foxes, or (3), which says that the fact that all vixens are foxes plays some role in explaining the truth of S. Our question, then, is which of these premises (if any) is a suitable target. And at least one theorist has recently responded to the objection from worldly fact by suggesting that there’s reason for conventionalists to be suspicious of (3).

That premise, recall, is just a consequence of the standard view of the relationship between the truth of S and the fact that all vixens are foxes, according to which the latter fact and the fact that S says that all vixens are foxes, taken together, fully explain the truth of S. But Warren (2015b: 91) argues that conventionalists have reason to think this standard view gets the order of explanation backward: conventionalist views “are most

31 The weaker premise can be stated as follows: there’s no fact about convention that v (i.e., the fact that all vixens are foxes) plays a role in explaining and that itself plays a role in explaining the truth of S. Or, more formally:

\[(4') \forall f(f \in C \rightarrow -S \text{ is true partly in virtue of } f \land f \text{ obtains partly in virtue of all vixens' being foxes})\]

And we can see why conventionalists are committed to this premise by thinking again about the Naturalist-Friendliness constraint. The reason conventionalism has any hope of meeting that constraint in the first place, after all, is that facts about convention are supposed to be facts our reliability about which isn’t mysterious by naturalist lights. But v is a fact about how things are with vixens, a fact our reliability about which is mysterious. So any facts that obtain even partly in virtue of v are also going to be facts our reliability about which is mysterious. And that means that, if conventionalism is to do the epistemological work it’s intended to do, v can’t have any role to play in explaining facts about convention. That is, conventionalists are committed to the following claim:

\[(4'') \forall f(f \in C \rightarrow -\lnot f \text{ obtains partly in virtue of all vixens' being foxes})\]

So, since (4'') is strictly weaker than (4'), conventionalists are committed to (4'') as well.

As for why the argument remains valid when (4) is replaced with (4''): Recall that the point of adding (4) as a premise was to ensure that there’s some fact not about convention in \(\Gamma\). (4) does this by ruling out the possibility that \(\Gamma\) contains some fact f such that v plays a role in explaining f and f plays a role in explaining the truth of S. But even if this possibility can’t be ruled out, there’s no problem for the objection from worldly fact as long as no such f is a fact about convention—i.e., as long as (4') is true—for the simple reason that, if \(\Gamma\) does contain any such f, it thereby contains a fact not about convention.
naturally launched against the background of meta-semantic theories that make dispositions to use sentences central and take any explanatory work done by propositions [and, presumably, by worldly (i.e., nonlinguistic) facts] to be less fundamental. Or, to put Warren's point another way: While conventionalists need not be skeptical of the existence of propositions, worldly facts, and the like, they should deny that such entities have any role to play in explaining the truth of sentences that are true by convention. After all, linguistic conventions plausibly are conventions for the use of expressions, and it seems odd to suppose that, even when our conventions for the use of a sentence meet whatever conditions they need to meet to provide for a full explanation of the truth of that sentence, that explanation must nevertheless take a detour out of the realm of language and into the realm of worldly facts (or propositions). One might have thought an explanation of one linguistic phenomenon in terms of another wouldn't need to be so indirect. If this is right, conventionalists have good reason to reject (3) even independently of its role in the objection from worldly fact.

Though Warren's claims have a certain plausibility, it seems to me that one of the central theoretical roles of facts (and propositions) is to explain the truth values of sentences. So I worry that, insofar as we accept the existence of worldly facts (or propositions) at all, we may be committed to giving them a role to play in explaining the truth values of even those sentences that are true by convention. In any case, though, we need not discuss this matter further, simply because there's a more fundamental difficulty here: rejecting (3)—whether there's independent reason to do so or not—turns out to be of no help to conventionalists. To see why this is, though, will require thinking a bit more about conventionalists' epistemological commitments.

Recall that, by Object-Level Belief, conventionalists owe an explanation of our ability to reliably form true object-level beliefs about facts such as the fact that all vixens are foxes. And their general strategy is (of course) to explain our reliability about these sorts of nonempirical matters by appeal to linguistic conventions. Presumably, then, what they'll need to do here is to appeal to our conventions for the use of $S$ and then to say something about what the connection is supposed to be between those conventions and the fact that all vixens are foxes. (A connection of the right sort will, one hopes, be able to do the epistemological work of explaining, in a way that respects Naturalist-Friendliness, our access to that fact.) But what exactly should they say?

One story to tell here, the standard story, is that the connection is as follows: the conventions for the use of $S$ and the fact that all vixens are foxes, taken together, fully explain the truth of $S$. To endorse (3) is to embrace this story, and the upshot of the objection from worldly fact (as I've reconstructed it) is that—unless the fact that all vixens are foxes itself obtains by convention—this story is inconsistent with conventionalism.

Warren's point, though, is that conventionalists can (and perhaps should) respond to the objection by denying that this story is mandatory. And this is fair enough—perhaps it's true that opponents of conventionalism are illegitimately relying on a contentious metasemantic theory. But we need to remember that conventionalists still owe some story here. They need to say something about the connection between our conventions for the use of $S$ and the fact that all vixens are foxes, and the story they tell must give them the resources to explain our access to this latter fact. Furthermore, Naturalist-Friendliness places strict constraints on what the story may look like—it tells us that the resulting explanation of our access to the fact that all vixens are foxes must not re-
quire appeal to any facts our access to which remains mysterious.

The question, then, is what sort of story, if any, remains available. And the answer, I claim, is that there's no story for conventionalists to tell here. After all, if they deny that the fact that all vixens are foxes—i.e., \(v\)—plays a role in explaining the truth of \(S\), only three possibilities remain: (i) the truth of \(S\) fully explains \(v\); (ii) the truth of \(S\) plays a role in explaining but doesn't fully explain \(v\); or (iii) neither plays a role in explaining the other. And conventionalists, it turns out, can't accept any of these. I consider them in turn.

In a certain sense, (i) is the most promising—it ensures that there's a tight explanatory relationship between \(v\) and the truth of \(S\), a relationship that can explain how our access to the latter might give us access to the former. The problem with (i), though, is straightforward: it doesn't actually give conventionalists a way to avoid (C). And this is independent of epistemological considerations. To embrace the claim that the truth of \(S\) fully explains the fact that all vixens are foxes, after all, is just to replace (3) with the following:

\[(3')\] All vixens are foxes purely in virtue of \(S\)’s being true

And the argument for (C) remains valid when (3) is replaced with (3'). In fact, (1) and (4) aren't even needed in this new version of the argument: (C) follows just from (2) and (3'). By (3'), \(v\) obtains purely in virtue of the truth of \(S\), which means that, if \(S\) itself were true purely in virtue of facts about convention, \(v\) would thereby also obtain purely in virtue of facts about convention. But (2) tells us that \(v\) doesn't obtain purely in virtue of facts about convention. By (2) and (3'), then, \(S\) can't be true purely in virtue of facts about convention, which is to say that (C) must be true.

So it's easy to show that, if conventionalists embrace (i), the objection from worldly fact loses none of its force. Showing why conventionalists can't embrace either (ii) or (iii), though, isn't so straightforward—neither of them leads to outright inconsistency with conventionalism. The problem is instead an epistemological one: neither (ii) nor (iii) can give conventionalists the resources to explain, in a naturalist-friendly way, our near-perfect reliability about facts like the fact that all vixens are foxes. And this is just to say that any conventionalist view that embraces either claim will thereby fail to do the epistemological work conventionalism is intended to do.

We can see why this is by noting that, if either claim is correct, then conventionalists have no way of explaining why \(S\)'s being true by convention is sufficient to guarantee that \(v\) obtains. Let's suppose that either (ii) or (iii) is true: either the truth of \(S\) plays some role in explaining \(v\) but doesn't fully explain it, or neither plays any role in explaining the other. In either case, there are some facts that play a role in explaining \(v\) but that aren't explained by the truth of \(S\). Furthermore, at least one of these facts—call it \(f\)—can't either be a fact about convention or obtain purely in virtue of facts about convention (nor can its role in explaining \(v\) be derivative of its role in explaining any fact about convention), since (2) rules out the possibility that \(v\) obtains purely in virtue of convention.

---

32 Assuming, again, that the purely-in-virtue-of relation is transitive.

33 Unless, of course, \(v\) is fundamental. But if that's so, \(S\)'s being true by convention certainly can't guarantee that \(v\) obtains—whether \(v\) obtains, in that case, has nothing whatsoever to do with what conventions are in place.
So, if conventionalists are to explain our near-perfect reliability about facts like v (and thereby meet the Object-Level Belief constraint), they must explain how we can be sure, in cases like this, that some suitable f obtains. Otherwise, they won’t be able to explain our ability to rule out the following possibility: that v fails to obtain as a result of there being no suitable f available. But conventionalism just doesn’t give us the resources to explain, in any naturalist-friendly way, how we have access to any such f.\(^3\) And that means that conventionalists who go in for either (ii) or (iii) will thereby be unable to meet the Naturalist-Friendliness constraint.

Summing up: conventionalists can’t avoid (C) by replacing (3) with either the claim that v obtains only partly in virtue of the truth of S or the claim that neither obtains even partly in virtue of the other. To try to do so, after all, would be to abandon the central task for which conventionalism was designed: to provide a naturalist-friendly explanation of our ability to reliably form true object-level beliefs in certain domains. And we’ve already seen that conventionalists also can’t avoid (C) by replacing (3) with the claim that v obtains purely in virtue of the truth of S, since the argument for (C) remains valid when we replace (3) with (3’). So, since the only way to reject (3) is to embrace one of those three claims, there’s no way for conventionalists to avoid (C) just by rejecting (3).

What this tells us is that conventionalists’ only remaining option is to reject (2)—doing so is the only way they can resist (C) and thereby hold on to their view. That is, those who want to affirm that facts about convention fully explain the truth of the sentence “All vixens are foxes” are thereby also committed to affirming that facts about convention fully explain the fact that all vixens are foxes. To maintain their view, then, conventionalists must find a way to make this latter claim plausible.

This, then, is my primary conclusion: if conventionalism is to remain a live option, conventionalists need to provide a plausible metasemantic theory according to which certain facts about the world outside of us (such as the fact that all vixens are foxes) obtain by convention alone.

Opponents of conventionalism will, of course, welcome this conclusion. After all,\(^3\)

---

\(^3\)I want to note that there’s a gap in this argument. We haven’t ruled out the possibility that the relevant f plays some role in explaining v and also, independently, plays some additional role in explaining some fact about convention c (i.e., that f’s role in explaining v isn’t merely derivative of its role in explaining c). If this were true, it might be possible to tell some story of the following form: Since c is a fact about what conventions are in place, it’s no mystery how we have access to c. But facts about convention aren’t fundamental—they obtain in virtue of certain other facts about us, facts about our linguistic dispositions, brain states, etc. Our access to c, then, puts us in a position to be sure also that some other facts obtain (namely, the facts, whatever they are, in virtue of which c obtains). So, if we’re somehow in a position to be sure that, whenever c obtains, one of the facts in virtue of which it obtains will also, independently, play a role in explaining v, then we may be able to be sure, just by knowing what conventions are in place, that some suitable f will be available.

But there at least two problems with this strategy. First, no plausible story of this sort has ever been given, and it’s not at all clear what a plausible story might look like. And second (and more importantly), any motivation for endorsing (2) is equally a motivation for ruling out the possibility that f has independent roles to play in explaining v and c. The central reason for endorsing (2), after all, is to respect the thought that facts about how things are with us can’t fully explain facts about how things are with the world outside of us, such as the fact that all vixens are foxes. But if f plays a role in explaining c, then f is a fact about how things are with us, despite the fact that it’s not itself a fact about convention. So, if conventionalists embrace this strategy, they’re still forced to say that v obtains purely in virtue of facts about us. And since the motivation for endorsing (2) in the first place is to avoid that result, conventionalists who remain committed to (2) shouldn’t embrace this strategy.
the very idea that facts about the world might obtain by convention is widely believed to be obviously absurd, and if this belief is correct, conventionalism is false: we've just seen that the only way to avoid (C) is to deny (2), which means that, if denying (2) is absurd, then (C) is unavoidable.

But we should be careful: the argument here is conclusive against conventionalism only if it is indeed obvious that facts about the world can't obtain by convention. Our question, then, is what reasons conventionalism's opponents can give for taking this latter claim to be obvious. (Surely the obviousness isn't supposed to just be intuitive—one of our reasons for taking conventionalism seriously in the first place, after all, is that there are concerns about the epistemic status of mere appeals to intuition.) And the answer usually given is this: the claim that facts about the world can be a matter of convention is obviously false for the simple reason that it has absurd implications, implications that are inconsistent with certain of the things we take to be a priori. Conventionalists, then, have a way forward here only if it can be shown that their doctrine doesn't in fact have the implications in question.

As it turns out, this work can be done. In fact, it has already been done— theorists sympathetic to conventionalism have shown that the doctrine doesn't have the absurd implications it's often taken to have. But these theorists' findings have, I think, gone underappreciated, and so I close with a brief discussion of what the absurd implications of denying (2) are purported to be and what conventionalists can say about those purported implications.

5 Two objections: Idealism and contingency

Those who endorse the claim that facts about convention can't fully explain the fact that all vixens are foxes—i.e., those who endorse (2)—aren't always explicit about their reasons, but as far as I can tell, these theorists tend to be motivated by one or both of the following sorts of worry: that denying (2) would commit us to idealism and that denying it would commit us to incorrect verdicts about the truth values of certain modal claims. So, if we can show that the denial of (2) doesn't in fact carry with it these commitments, we can thereby show that conventionalism remains worth exploring.

The first thing to note here is that the basic thought underlying both sorts of worry is the following: If the fact that all vixens are foxes obtains purely in virtue of facts about what conventions are in place, then our conventions somehow determine whether all vixens are foxes, which means that whether all vixens are foxes will be in some way sensitive to whether the relevant facts about convention obtain. That is, whether all vixens are foxes is going to vary depending on what conventions are in place. This sort of convention-sensitivity is what's thought to be at the root of both of the problematic commitments mentioned above. I discuss them separately.

It's easy enough to see how convention-sensitivity gives rise to the worry about modal claims. Facts about what conventions are in place, after all, are contingent: our linguistic conventions could easily have been different. So, if it's really the case that whether a given fact f obtains varies depending on whether certain facts about convention obtain, it's natural to infer that, in at least some of the nearby possible worlds where the relevant facts about convention don't obtain, f doesn't obtain either. And if that's right, then
modal claims such as the following turn out to be true:

- It's contingent that all vixens are foxes.\footnote{One theorist who takes conventionalists to be committed to claims of this sort is Blackburn (1986: 121)—he argues that necessities can't be explained by contingent facts about convention, since "the explanation, if good, would undermine the original modal status: if that's all there is to it, then [for example] twice two does not have to be four."}
- Necessarily, if we accept the sentence "All vixens are female goats" as a matter of convention, then all vixens are female goats.\footnote{This claim is a variant of the one that's the basis of Lewy's (1976) central objection to conventionalism, as discussed in §1 above.}
- If, as a matter of convention, we applied the word 'vixen' to all and only those things to which we were willing to apply both 'female' and 'goat', not all vixens would be foxes.\footnote{Boghossian (1996) and Stroud (1984) both cite conventionalists' purported commitment to counterfactuals like this one as a serious problem for conventionalism, as does C. I. Lewis (1946: 148), whose argument against truth by convention goes via the following claim: "If the conventions were otherwise, the manner of telling would be different, but what is to be told, and the truth or falsity of it, would remain the same."}

And so on. The problem, of course, is that each of these modal dependence claims seems obviously false—the supposition that whether all vixens are foxes is sensitive to contingent facts about convention appears to have led us into absurdity.

As defenders of conventionalism have pointed out, though, this sort of worry is based on a misunderstanding. What conventionalists are committed to, again, is the claim that our conventions determine whether certain facts obtain, so that whether these facts obtain varies depending on what conventions are in place. And it's tempting, to be sure, to conclude that conventionalists are thereby committed to the claim that different facts would have obtained had our conventions been different—it's natural to suppose, after all, that the kind of covariance that's in play here just entails the truth of modal dependence claims of this sort. But that supposition, reasonable as it may seem, turns out to be mistaken.

It's true enough that, in the usual cases where one fact determines another in such a way that we can change whether the second obtains by changing whether the first obtains (such as cases of causal dependence), we can read off the dependence from the facts' modal profiles. But given the special role conventions play in our linguistic practice, a straightforward explanation is available of why, in the particular case of worldly facts that vary depending on whether certain facts about convention obtain, we should expect the dependence not to be reflected in the modal relationships among these facts. Observe, first of all, that our conventions have an important role to play even in our descriptions of possible worlds other than our own: when we talk about what's the case in possible but nonactual situations, we do so using our own language, with all its attendant conventions. Furthermore, one of our conventions, plausibly, is to be willing to apply the predicate 'vixen' to anything to which we're willing to apply both 'female' and 'fox'—regardless of whether the thing we're talking about exists in the actual world or in some other possible situation in which our conventions (i.e., the conventions of our counterparts) are different.\footnote{As Wright (1985: 190) puts it, our talk is governed by the general convention that "what it is true to say}
foxes doesn't modally depend on our conventions, and the explanation has nothing to do with whether conventionalism is correct—it simply has to do with what our conventions direct us to say about possible worlds other than our own. (Consider: there's nothing obviously incoherent about a language in which the convention for describing what's the case in a possible but nonactual situation in which our linguistic conventions are different is to defer to the linguistic conventions that are in place among our counterparts in that situation.\textsuperscript{39} We don't in fact speak such a language, but we very well could have. The fact that we've chosen \textit{not} to speak such a language is surely irrelevant to whether conventionalism is correct.)

To put the point slightly differently: The reason we tend to understand the sort of covariance that's in play here in terms of modal dependence is that thinking about what's the case in possible but nonactual situations is our usual method for getting a grip on how facts are related to one another—we learn about the determination relations among facts by imagining worlds where certain facts don't obtain and so can't participate in any such relations. But in the particular case of facts about convention, this method is ineffective. Given the way our conventions work, thinking about possible worlds with different conventions is \textit{not} a way of preventing our actual conventions from playing a determinative role: again, even when we talk about possible worlds other than our own, our talk is governed by our actual conventions. What this suggests is that thinking about what's the case in possible but nonactual situations just isn't the right way to settle the question of whether the fact that (e.g.) all vixens are foxes obtains purely in virtue of facts about convention.

Of course, this isn't the end of the discussion. It's still the case that the usual way of making sense of the sort of covariance that's in play here is in terms of modal dependence. Conventionalists, then, need an alternative way of understanding the claim that whether a given fact obtains varies depending on what conventions are in place, one on which that claim can be true even if all the usual modal dependence claims are false. That is, they need to say just what it is for a fact to be convention-sensitive, if convention-sensitivity doesn't entail modal dependence. What (if anything) can they say here?\textsuperscript{40}

Fortunately, a satisfying answer is available. The motivation for conventionalism, remember, is epistemological: the goal is to provide a naturalist-friendly explanation of how we manage to have nonobservational access to facts in certain domains. And as we've just seen, conventionalists, in order to do this explanatory work, must provide a story on which the facts in question obtain purely in virtue of our linguistic conventions.

Of a hypothetical state of affairs... is to be determined by reference to our actual linguistic conventions, even if those are not the conventions that would then obtain. And Sidelle (1989: 7) explains, similarly, that "if it is a convention of ours that nothing in any possible situation counts as water if it is not composed of H\textsubscript{2}O, then this very convention tells us that in the subclass of possible situations in which we have different conventions, still, nothing counts as water that is not H\textsubscript{2}O: that is, that it is necessary that water is H\textsubscript{2}O".\textsuperscript{39}

\textsuperscript{39} The most explicit statement of this point comes from Wright (1985: 191).

\textsuperscript{40} Sider (2001: 102) expresses a version of this worry, saying that it's "unclear just what sort of dependence of truth upon conventions is supposed to be distinctive of conventionalism" given that "the conventionalist will surely deny counterfactual or temporal dependence". The worry also appears to be the basis of Elder's case against conventionalism—he suggests that denying that counterfactual dependence follows from a fact's obtaining by convention would be "a desperate move indeed" (2006: 14).
when facts obtain purely in virtue of convention, there’s no mystery about how we have nonobservational access to those facts. And for that to be so, it must be that our conventions can’t lead us astray: the relevant facts must somehow be guaranteed to obtain given that the associated conventions are in place. So, for any sentence we accept as a matter of convention, what that sentence says is going to be true—there’s no way for our conventions to be mistaken.41

This, then, is the sense in which the fact that all vixens are foxes is convention-sensitive, according to conventionalism: given that the relevant conventions governing our use of the sentence “All vixens are foxes” are in place, the truth of that sentence is guaranteed—as is the obtaining of a corresponding fact. And on this way of understanding convention-sensitivity, conventionalists’ commitment to the convention-sensitivity of the fact that all vixens are foxes doesn’t commit them to any of the above modal dependence claims. They can instead endorse the usual picture about possible but nonactual situations in which our conventions are different: that, in a possible situation in which (say) we applied the word ‘vixen’ to all and only those things to which we were also willing to apply both ‘female’ and ‘goat’, the sentence “All vixens are goats” would be true, despite the fact that all vixens would still be foxes (and so wouldn’t be goats), since the sentence would then mean something different than it does in our actual language. What conventionalists add is only that the sentence “All vixens are goats” is guaranteed to be true in that situation, in the sense that no cooperation from the convention-independent world is required—the truth of the sentence is a matter of convention alone and so isn’t hostage to whether some suitable corresponding fact is antecedently available. And it’s clear enough that the addition of this claim doesn’t render the picture incoherent.

The worry about modal claims, then, has been answered. As it turns out, the claim that the fact that all vixens are foxes is convention-sensitive entails none of the obviously false modal dependence claims it’s been taken to entail.

But thinking about the sort of convention-sensitivity that’s in play here brings us back to the other worry mentioned at the beginning of this section: that conventionalists’ commitment to convention-sensitivity might bring along with it a commitment to some absurd sort of idealism. After all, if the facts really do vary with convention in such a way that our conventions can’t be mistaken, it’s natural to try to explain this phenomenon by claiming that reality is mentally constructed. And that claim, historical pedigree notwithstanding, seems absurd.42

Here again, though, a response is available, and we can again get a grip on it by remembering that the epistemological purpose of conventionalism is to explain how we manage to access certain facts our access to which is otherwise mysterious. This means that, in order for conventionalism to have any epistemological work to do, there must exist some such facts. That is, there must be domains of facts such that questions about whether those facts obtain can’t be settled by what’s given to us via observation (or by any other means, though conventionalists tend to be naturalists and empiricists and so tend not to countenance other means of access). After all, the only domains about which

---

41 This is a version of a point made by Sidelle (2009: 234–235).
42 Explicit expressions of this worry in the literature are rare. (Stroud’s (1984: chap. 5) discussion of Carnap’s conventionalism is an exception, although Stroud seems to take the worry about idealism to be equivalent to the worry about modal claims.) Still, I’ve heard it expressed fairly often in conversation, and so I think it’s worth saying something about.
there's any reason to be a conventionalist are those domains for which we don't otherwise have a satisfying epistemology—i.e., those about which observation is neutral (as is any other means of access to the world we take ourselves to have). So, on any reasonable conventionalist view, the only facts that can be fully explained by our conventions are facts in these epistemologically problematic domains.

Even according to conventionalists, then, the power of conventions is limited: only certain sorts of questions about what the world is like—namely, those that we don't otherwise have any good way of answering—may have the answers they do as a matter of convention. (Exactly which questions these are is up for debate, of course; which domains of facts are amenable to a conventionalist treatment will depend on just how much of our access to the world is unproblematic—that is, on just how much we can take to be given to us, by observation or any other means. Determining the limits of a reasonable conventionalism, then, will require serious epistemological investigation. But we need not engage in any such investigation here—our discussion can remain schematic.) And the questions we do have a good way of answering may very well have answers that have nothing to do with our mental states, which means there remains room for a mind-independent reality. According to Sidelle, for instance, his modal conventionalism isn't "in any interesting way idealist" despite the fact that he takes the modal properties (and hence the boundaries) of objects to be a matter of convention, since "the 'material' [out of which we carve objects via our conventions] is taken to be wholly mind-independent" (2010: 109).

And if we abstract away from Sidelle's particular brand of conventionalism, we can see more generally how a conventionalist picture is going to work: "The world", as Einheuser (2006: 461) puts it, "provides some material, the substratum (or stuff), which is neutral with respect to the features that are taken to be conventional", and "onto this substratum, features of the kind in question can be conventionally imposed in many different ways". The substratum, in other words, consists of what's given to us via observation (or by some other unproblematic means of access), and we, in adopting the relevant conventions, lay some additional structure over the top of this substratum. And insofar as worldly structure is conventional in this way, facts about how the world is structured obtain purely in virtue of convention, since they obtain regardless of what the substratum is like.

Now, exactly how much of the world's structure is taken to be conventional in this way will depend on the limits of the particular conventionalist theory we're talking about—our conventions may, for instance, be responsible for the modal properties of material objects, or for all properties of material objects, or for the entire realm of abstract objects, or for all of the above and more. But again, we need not settle the details here. For our purposes, the point is just that there's a substratum, and it's distinct from whatever is conventional. And this means that the sort of convention-sensitivity to which conventionalists are committed does not entail that reality is mentally constructed: even if certain of the facts of the world obtain purely in virtue of convention

---

43 Einheuser uses this idea to develop a conventionalist-friendly generalization of possible world semantics in which a world is understood as a pair containing a substratum and a “carving” (roughly, a function from substrata to arrangements of whatever features are taken to be conventional), and then she uses this formal framework to give a rigorous argument showing that convention-sensitivity doesn’t entail counterfactual dependence.
(and so are convention-sensitive), the possibility remains that there exists an entirely mind-independent realm (i.e., the substratum) and that what explains why the facts in question obtain is just that we have, by convention, overlaid some additional structure on this mind-independent realm.

What all this shows is that conventionalists’ commitment to the claim that facts about convention fully explain the fact that all vixens are foxes—and to the associated claim that whether this latter fact obtains is sensitive to whether the relevant facts about convention obtain—doesn’t in fact carry with it either of the commitments that have motivated opponents of conventionalism to reject those claims (i.e., to endorse (2)). So the quick dismissals that are common in the literature are all of them premature. Conventionalism stands unfutred. I suggest, then, that, given its unique epistemological advantages, conventionalism remains worth exploring.

6 Conclusion

We’ve covered quite a bit of ground here. We began by asking how, exactly, the conventionalist doctrine is to be understood, and though we didn’t arrive at a fully precise answer, we were able to gain some clarity as to what’s distinctive about conventionalism by thinking about its epistemological motivations. Then we provided a new rendering of the most influential argument against conventionalism, the objection from worldly fact, and in so doing we made transparent how that argument works and what assumptions it requires. Then, because this new rendering made possible a closer examination of the argument, we were able to show conclusively that conventionalists’ only viable strategy for maintaining their view in the face of that argument is to embrace the claim that facts about our linguistic conventions can fully explain (e.g.) the fact that all vixens are foxes—that is, the claim that there’s a sense in which our conventions have the power to make the world one way rather than another. Finally, we showed that this claim, though unorthodox, doesn’t in fact have the absurd implications it’s often taken to have, which means conventionalism can’t be dismissed out of hand: there remains a promising way forward for the conventionalist project. Conventionalism, then, remains worthy of serious investigation, especially given its unique epistemological promise.

But this isn’t to say that our work is done. So far, we’ve been able to show that the claim that facts about convention can fully explain facts about the world isn’t obviously absurd, which is a start. But to demonstrate that conventionalism is really a tenable view, we need to do much more: we need to provide a plausible metasemantic theory on which that claim is true. That, though, is a task for another time.44

References


44 Thanks to Zachary Barnett, David Christensen, Nina Emery, Tobias Fuchs, Phillip Galligan, Geoffrey Grossman, Louis Gularte, Yongming Han, Richard Kimberly Heck, Christopher Hill, Paul Horwich, Iain Laidley, Rachel Leadon, Miquel Miralbés del Pino, Mary Renaud, Kirun Sankaran, Joshua Schechter, and Leo Yan for helpful discussion of the material in this paper.


