4-2015

What We Talk About When We Talk About Content Externalism

Jeff Engelhardt
Dickinson College

Follow this and additional works at: http://scholar.dickinson.edu/faculty_publications
Part of the Philosophy Commons

Recommended Citation
Engelhardt, Jeff, "What We Talk About When We Talk About Content Externalism" (2015). Dickinson College Faculty Publications. Paper 149.
http://scholar.dickinson.edu/faculty_publications/149

This article is brought to you for free and open access by Dickinson Scholar. It has been accepted for inclusion by an authorized administrator. For more information, please contact scholar@dickinson.edu.
What we talk about when we talk about content externalism

Abstract: Some Content Externalists claim that if C is a theoretical concept and “C” expresses C, then the content of C in a community at a time is determined by how some members of the community at the time—call them “experts”—understand C or use “C”. Thus, when non-expert Chauncey utters “C”, the content of the concept he expresses does not depend entirely on his intrinsic physical properties, contra the claims of Content Internalism. This paper proposes that “concept” expresses a theoretical concept, such that the externalist’s insights should apply to how we understand claims expressing the view itself and to how we evaluate the arguments alleged to motivate it. With respect to the first, I argue that the content externalist (a) should regard it as unclear at present which proposition her theory expresses, and (b) should take it that content externalism teaches us about our linguistic community rather than about the metaphysical nature of concepts. With respect to the second, I argue that by externalism’s own lights, the famous externalist thought experiments shouldn’t establish content externalism. In conclusion, I suggest that making sense of content externalism requires presupposing internalism.

Keywords: Content Externalism; theoretical concepts; Content Internalism; concepts; Tyler Burge.

Words: 8,804

1. Introduction

This paper concerns Tyler Burge’s Content Externalism. According to Burge, words have their referents determined by social features of one’s linguistic community. In Burge’s famous example, it is socially-determined whether an individual’s utterances of “arthritis” refer to arthritis or something else, tharthritis. If this holds as generally as Burge says it does, then it should also be socially-determined whether an individual’s utterances of “concept” refer to concepts or something else. If so, then utterances of, say, “concepts ain’t all in the head” express different propositions in different linguistic communities. Content externalism should hold for attempted expressions of content externalism.
If Burge’s claim ‘applies to itself’, then three noteworthy consequences follow. First, in order to determine which proposition Burge’s claim expresses, we must determine what content is expressed by “concept” as used in our linguistic community; if we haven’t done so, then we don’t know, according to content externalism itself, what content externalists claim. Second, if we learn that Burge’s claim does express a true proposition, all we can glean from it is that the content we express with “concept” comports well with externalist ideas; this is a contingent sociological fact. Third, Burge’s claim calls for justification by empirical inquiry, not solely by thought experiments. Since the claim expresses a true proposition just in case our concept comports with externalism, it needs to be shown what our concept happens to be. This calls for the methods of sociology, not conceptual analysis.

Here’s the plan for the paper. In the next two sections (2-3), I’ll sketch Burge’s reasoning for the claim that concerns us and offer reasons for thinking that it applies to itself. In the following three sections (4-6), I’ll argue for each of the consequences sketched in the preceding paragraph. Section 4 concludes with a puzzle for content externalism. Section 7 considers a reply. As a ‘moral’ of the story, I suggest that the externalist presupposes content internalism in order to articulate, understand, and argue for her position. But this is not to say that content externalism must be given up. The conclusion sketches a view that could resist most of the difficulties raised here.

But first, a few clarifications. In talking about concepts here, I’ll just take them to be the constituents of thought; I’ll often talk as though they’re mental representations, but I don’t mean to presume that they’re irreducibly representational. If representations ‘reduce’ to something else, the argument should still go through. I take it to be a standing question whether concepts
supervene on an individual’s intrinsic physical properties or not. Indeed, it’s one of the questions this paper is asking. As such, I shouldn’t assume what they supervene on in advance. I’ll be neutral so far as I can. Unless otherwise noted, then, when I say “concepts” in what follows, I’m referring to the contentful constituents of thought that have their contents determined however we have our conceptual contents determined. And except where I say otherwise, I’ll take our concept CONCEPT\(^1\) and our word “concept” to refer to those very mental phenomena.

I’ll contrast Burge’s externalism with content internalism. Internalism claims that the contents of an individual’s concepts are fully determined by the individual’s intrinsic properties (and usually, the individual’s intrinsic physical properties). For our purposes, it should suffice to take intrinsic properties to be those properties such that they never differ between duplicates. I’ll take it that if Oscar and Twin Oscar are physical duplicates and F is an intrinsic physical property, Oscar is F iff Twin Oscar is too. The problems with this characterization of intrinsicality shouldn’t disturb the argument posed here.

2. Content externalism

For the sake of having a claim to focus on, I propose the following formulation of Burge’s claim:

[CE] If C is a theoretical concept and “C” expresses C in a community at a time, then the content of C is determined by intellectual norms and practices surrounding C and “C” in the community at the time. (See, e.g. Burge 1979, 1982, 1986a, 1986b, 1989)

\(^1\) I use small caps (e.g. CONCEPT) to indicate that I’m referring to the concept itself and not its referent.
The intellectual norms and practices that determine the content of a concept in a community are complex and multifarious. Indeed, Burge tells us that “…we are not in a position to provide a very rich and detailed account of the nexus of abilities that go into having a concept…” (Burge 2003, 518) Nonetheless, we can get a glimpse of the ways norms influence language use.

An appreciation of the ways that people depend on one another in language use derives from experience of specializations, differences of positioning, and differences in background knowledge among people within a culture. (Burge 2006, 169)

The norms concern appropriate use of terms, conversational deference, and practices of negotiating disagreements about a term’s referent; they turn in part on who does and doesn’t ‘count’ as an intellectual authority, how authorities with respect to a particular concept are regarded, how informed non-authorities are about some concept; and so on. (See Burge 1986b)

The view has it, then, that concepts are determined by social interactions of various sorts, but we presently can’t provide much detail about the abilities that underwrite concept possession. I’ll return to these points in section 4, but let me note that I’m not focusing on this point for criticism. Most of us would agree that our present understanding of concepts is far from complete and detailed. It isn’t a problem peculiar to Burge’s view; it’s simply where our inquiry into concepts stands at present.

In what follows, it will sometimes be useful to refer to the individuals and practices in a linguistic community rather than simply to a community’s norms. For ease of expression, I’ll refer to those with the intellectual authority in a community to enforce and question norms regarding some concept C as “the experts”; all others are “non-experts”. This isn’t to presume that there are always clear and explicitly recognized authorities on a given concept or that those
with less overall intellectual authority never influence a given word’s use or a given concept’s content. It’s just for ease of expression and to help us imagine the differences between two communities that would underwrite a difference in the concepts available to them.

Burge makes his case for content externalism primarily by appeal to a thought experiment, the tharthritis thought experiment. (See, for example, Burge 1979, 1982, and 1986a) Imagine Ghassan, who has a belief that he would express by saying, “I have arthritis in my thigh”. Ghassan is unaware that as experts in our community use the term, “arthritis” refers to a rheumatoid ailment that occurs only in joints. It’s impossible to have arthritis in one’s thigh. Ghassan’s belief, then, is false. But what if we used the term differently? Burge asks us to imagine a counterfactual scenario in which Ghassan (or his counterpart) lives in a community identical to his actual community in every respect except that experts there use the word “arthritis” to refer to a rheumatoid ailment that occurs in joints but can also occur in one’s thigh. In this counterfactual case, when Ghassan says, “I have arthritis in my thigh”, then prima facie it’s possible that he expresses a true belief. If he has the ailment that his community refers to with the word “arthritis”, then what he says is true.

But if we suppose that Ghassan is physically the same in the actual and counterfactual cases, it would seem that if he has arthritis in one case, he must have it in the other as well. How can his utterances differ in truth value if he doesn’t change physically? If he expresses the same proposition representing the same physical body as having arthritis, then if one utterance

---

2 If this exception entails other differences between Ghassan’s actual and counterfactual communities, then those differences apply as well. If they entail differences between actual and counterfactual Ghassan, then we take it that these differences between actual and counterfactual Ghassan don’t explain the differences in their concepts. Rather, those differences are explained by (expert) usage in their communities.
expresses a truth, so should the other. The answer is that Ghassan’s utterances don’t express the same proposition. Why? Because Ghassan has access to different conceptual contents in the two cases. Actual Ghassan expresses a belief about arthritis, the rheumatoid ailment that our word “arthritis” refers to. Let’s say that the concept at play in his belief is ARTHRITIS. But Burge says that counterfactual Ghassan doesn’t have access to this concept—it’s not available in his community—so it’s unclear how he might have a belief about arthritis, much less express one. (See, e.g., Burge 1986a, 5; 1986b, 265; 1988, 650) Rather, he has to use the concepts that are available in his community. When he uses the word “arthritis”, he’s expressing a belief about a rheumatoid ailment that occurs in joints and thighs. Since arthritis can’t occur in thighs, counterfactual Ghassan’s belief doesn’t concern arthritis but some other ailment, call it “tharthritis”. The content of the concept at play in counterfactual Ghassan’s belief, then, must refer to tharthritis, and since the concept determines a different referent from the concept ARTHRITIS, it must have a different content. Indeed, it must be an entirely different concept: not ARTHRITIS but THARTHritis. This difference in their conceptual contents explains the intuition that one might speak truly while the other does not.

Given that actual Ghassan and counterfactual Ghassan are physical duplicates and yet their concepts differ, it must be that the content of a concept isn’t fully determined by an individual’s intrinsic physical properties. In this case, since the salient difference lies in the (expert) usage found in their communities, it seems that this difference accounts for the difference in concepts. Prima facie, each Ghassan’s concept derives its content from expert norms and practices in his linguistic community. Moreover, Burge thinks this point—that we depend on others to fix the contents of our concepts—applies to theoretical concepts like GENE, CANCER,
ARTHRITE, and CARBURETOR, observational concepts including RAIN, PENGUIN, and TIGER and to any concept with a public referent.

Such arguments go through for observational and theoretical notions, for percepts as well as concepts, for natural kind and non-natural kind notions, for notions that are the special preserve of experts...[and for] any notion that applies to public types of objects, properties, or events that are typically known by empirical means. (Burge 1986a, 6; see also 1989, 287-9)

Indeed, Burge claims that similar arguments apply to “any substantive notion that applies to physical objects, events, stuffs, properties...the notions may be as ordinary and as observational as one likes.” (Burge, 1986b, 264) And he thinks these points are especially obvious when it comes to non-observational terms.

The main points of the argument for social dependence apply equally, indeed even more obviously, to terms that are less closely associated with direct perception. (Burge 1989, 288)

Moreover, he thinks that his arguments leave no propositional attitude untouched.

[These arguments] suggest that virtually no propositional attitudes can be explicated in individualistic terms. Since the intentional notions in terms of which propositional attitudes are described are irreducibly non-individualistic, no purely individualistic account of these notions can possibly be adequate. (Burge 1982, 99)

As Burge sees it, each individual depends on other members of her linguistic community to fix the referents of her words; the referents then fix the contents of her concepts. Since propositional attitudes comprise concepts, which proposition any individual believes is partly determined, according to Burge, by features of her community.

[The individuation of a given individual’s mental contents depends partly on the nature (or what his fellows think to be the nature) of entities about which he or his fellows have de re beliefs. (Ibid 90, emphasis in original)
To know and explicate what a person believes *de dicto*, one must typically know something about what he believes *de re*, about what his fellows believe *de re* (and *de dicto*), about what entities they ostend, about what he and his fellows’ words mean, and about what entities fall in the extensions of their terms. (Ibid 95)

3. Externalism applies to CE

If content externalism applies to all concepts and CE expresses concepts, then externalism applies to the concepts that CE expresses. But I’d like to focus on one of these, namely, the concept expressed by “concept”. And, rather than assume that externalism applies to this concept, I’ll argue for it. It should be clear that one can’t reject the difficulties I raise here by claiming that the concept expressed by “concept” is one of the few that isn’t socially-determined. I’ll argue in this section, then, that “concept” expresses a theoretical concept. I take it that the theoretical concepts (to be clarified below) form a subset of the concepts such that “theoretical concept” refers to a subset of those entities, whichever they are, that we refer to with “concept”. If this is right, then content externalism concerns a subset of those things to which “concept” refers. To what does “concept” refer? If CONCEPT is a theoretical concept, then according to CE, the referent of the word “concept” and the content of the concept it expresses are both determined by intellectual norms and expert practices involving CONCEPT and “concept”. With this established, I’ll spell out an analog to the tharthritis thought experiment for “concept”.

What are theoretical concepts? I’ll follow Carnap and many others in taking them to be the concepts expressed by theoretical terms (or: terms in theoretical language); theoretical terms are those terms in a science, theory, or explanation that refer to unobservable events, properties, things, etc. The theoretical language, terms, and concepts contrast with the observation
language, its terms and concepts. Observation terms and concepts, of course, refer to observable events, properties, things, etc. As an example, Carnap proposes that terms referring to drive and potential in psychology are theoretical terms. (Carnap 1956, 38) I take it to be *prima facie* quite plausible that “concept” belongs to the theoretical language of both ‘everyday’ explanations of behavior and the explanations and theories offered in psychology, cognitive science, linguistics, and more.

Furthermore, the thought that CONCEPT and other concepts referring to mental phenomena are theoretical concepts has a fairly well-known philosophical history. Wilfrid Sellars asked us to imagine that terms referring to mental kinds were proposed by a mythical “Jones” as part of a theory that serves to explain and predict intelligent behavior.

Suppose, now, that in the attempt to account for the fact that his fellow men behave intelligently...Jones develops a theory according to which overt utterances are but the culmination of a process which begins with certain inner episodes. (Sellars 1963, 186; emphasis in original)

And after Sellars, many, many philosophers of mind took up the idea that mental concepts are theoretical concepts. (See, e.g. Fodor 1968 (e.g. xxi), 1987; Lewis 1966, 1970, 1972; Churchland 1981; Dennett 1987, 1991) I’ll take it as granted, then, that CONCEPT is a theoretical concept. The content externalist should accept, then, that the content of CONCEPT is socially determined.

Whereas it’s widely-accepted that some are experts on arthritis, things are far from clear with respect to concepts. The social structures involved in determining the referent of “concept” and the content of CONCEPT are less obvious and presumably more complex. But this alone needn’t cast any doubt on the concept’s social determination. *Prima facie*, not all members of our linguistic community have equal intellectual authority with respect to CONCEPT; Burge may say
that children and others just entering our language community count on the semantic division of labor to give their utterances of “concept” its meaning and reference. Although I won’t pretend to know what the norms and practices that determine the content of CONCEPT are, I will presume that there are some such norms and practices and that they hold as a matter of contingent fact.

Indeed, the assumptions noted above are taken just for the sake of concreteness in the examples I’ll use. If all concepts referring to mental phenomena are theoretical concepts, then we needn’t use CONCEPT. We may use any concept that refers to mental phenomena. Similarly, the arguments to follow don’t turn on the assumption that expert practices determine the content of CONCEPT. They turn, rather, on the assumption that some contingent feature or features of a linguistic community determine the content of CONCEPT. Accordingly, the arguments to follow still succeed if it turns out that the content of CONCEPT isn’t determined by expert use of “concept” but by children’s uses, by all and only whispered uses, etc.

Moreover, I refer to theoretical concepts here just in order to give grounds for thinking that Burge’s thesis applies to the concept CONCEPT. If it’s unclear what theoretical concepts are and whether CONCEPT is among them, there may be other grounds for claiming that Burge’s thesis applies to itself. After all, Burge thinks his social externalism applies to “any substantive notion that applies to physical objects, events, stuffs, properties…the notions may be as ordinary and as observational as one likes.” (Burge, 1986b, 264)

With these assumptions in place, we can give an analogue of the tharthritis thought experiment. If the content of the concept CONCEPT is determined by intellectual norms and social practices, then where norms and practices differ, the content of the concept should differ
as well. Just as the content of the concept expressed by “arthritis” may change from one community to the next, so may the content of the concept expressed by “concept”.

In demonstrating this, it is important to emphasize the parallels with Burge’s tharthritis thought experiment. I begin with the two contents that will be in play. Instead of ARTHRITIS and THARTHITIS, let me introduce ICONCEPT and ECONCEPT. Whereas both ARTHRITIS and THARTHITIS refer to rheumatoid ailments, both ICONCEPT and ECONCEPT refer to contentful mental phenomena. Where ARTHRITIS and THARTHITIS differ in that the latter refers to a rheumatoid condition that may also afflict thighs, ICONCEPT and ECONCEPT differ in the following way. ICONCEPT refers to contentful mental phenomena that have their contents fully determined by the intrinsic properties of the individuals who have them; ECONCEPT refers to contentful mental phenomena that have their contents partly determined by an individual’s intrinsic physical properties and partly by features of an individual’s linguistic community. In some cases, for instance, the content of an individual’s Econcept is determined by intellectual norms surrounding words that express the Econcept in question. ICONCEPT and ECONCEPT differ, then, in that the latter refers to mental phenomena that have their contents determined by intrinsic physical properties and something more. I hope the parallels with ARTHRITIS and THARTHITIS are clear. Note that I reserve CONCEPT for the concept we express with “concept”.

Each of the following is, of course, an open possibility: (i) CONCEPT = ICONCEPT, (ii) CONCEPT = ECONCEPT, (iii) CONCEPT ≠ ICONCEPT and CONCEPT ≠ ECONCEPT. We shouldn’t presume that

---

3 It may be, of course, that one or the other of these concepts (or both) fail to refer in actuality. I’m not supposing that either concept refers in actuality.
ICONCEPT or ECONCEPT is the concept that we express by “concept”. Keep in mind that both communities discussed below may be alien to us.

Just as actual Ghassan had the concept ARTHRITIS thanks to the actual community and counterfactual Ghassan had a different content thanks to his counterfactual community, we can so distribute ICONCEPT and ECONCEPT. Let IChauncey and EChauncey be non-actual counterparts. (As above, they’re non-actual because I don’t want to presume that our use of “concept” refers to either Iconcepts or Econcepts.) Both IChauncey and EChauncey speak and write a language that includes a word that they would write as “concept”. But, of course, their communities differ in the intellectual norms surrounding proper use of the word. Just as actual Ghassan’s local experts use “arthritis” to refer to a rheumatoid ailment that afflicts only joints, intellectual norms in IChauncey’s community make it so that “concept” refers to contentful mental phenomena that have their contents fully determined by an individual’s intrinsic physical properties. Just as counterfactual Ghassan’s local experts use “arthritis” to refer to a rheumatoid ailment that afflicts both joints and thighs, EChauncey’s local norms determine that “concept” refers to contentful mental phenomena that have their contents determined by both one’s intrinsic properties and some features of one’s social world. That is, IChauncey’s community uses “concept” to express ICONCEPT and to refer to Iconcepts; EChauncey’s community uses “concept” to express ECONCEPT and to refer to Econcepts. Taking the tharthritis thought experiment as our guide, we should expect, then, that when IChauncey says “concept”, he too expresses ICONCEPT and refers to Iconcepts; and of course the same goes, mutatis mutandis, for EChauncey, ECONCEPT, and Econcepts.
Suppose now that both Chaunceys say, “Concepts ain’t all in the head.” Since IChauncey refers to Iconcepts when he says “concept”, he’s claiming that Iconcepts aren’t all in the head; since EChauncey refers to Econcepts when he says “concept”, he’s claiming that Econcepts aren’t all in the head. They express different propositions. Prima facie, EChauncey’s claim is true, since Econcepts supervene on both an individual’s intrinsic properties and some properties extrinsic to the individual; IChauncey’s claim, however, may be false. Since Iconcepts supervene on an individual’s intrinsic physical properties, it might be that all of those properties are ‘in the head’ in such a way that it’s false to say that Iconcepts ain’t in the head. Suppose it is false: Iconcepts are all in the head. Just as actual Ghassan spoke falsely but counterfactual Ghassan didn’t, IChauncey is wrong but EChauncey isn’t.

We now have an example in which the content of the concept expressed by “concept” differs from one community to the next. Call it the “I/Econcepts thought experiment”. If content externalism is true for theoretical concepts and “concept” expresses a theoretical concept, then the I/Econcepts thought experiment carries just as much plausibility as the tharthritis thought experiment. More generally: if, according to content externalism, CONCEPT belongs to any set of concepts that have their contents socially determined, then some thought experiment like this one should be available.

In what follows, I’ll be using this example to develop the claims promised in the introduction, partly by asking whether CE, the formulation of Burge’s claim given above, refers to concepts or something else. That is, I’ll ask whether CE, a string of English words, successfully expresses content externalism or successfully refers to concepts. It is important, then, to distinguish between the inscription or utterance CE and content externalism, the theory
that CE seems to express. Accordingly, note well that “CE” in this paper refers to an utterance or inscription. I will use “content externalism” and “externalism” to refer to the proposition that CE seems to express, i.e. the proposition that conceptual contents are socially determined.

We may now make a preliminary, suggestive remark: if IChauncey utters CE, he will be wrong; but if EChauncey says it, he’ll be right. Since IChauncey’s referring to Iconcepts, and since Iconcepts do have their contents fully determined by an individual’s intrinsic physical properties, his claim is false. But since EChauncey is talking about Econcepts, his claim is true. And this is the case on the assumption that content externalism is true.

4. What does CE say?

Suppose Ghassan says, “Arthritis is common in my family.” Which proposition has he expressed? Has he said that arthritis is common in his family or that tharthritis is common in his family?

You don’t know. You don’t know which proposition he’s expressed because I didn’t specify which Ghassan was speaking. You don't know whether the experts in this Ghassan’s community determine that he expresses ARTHRITIS or THARTHITIS when he says “arthritis”. Although you have the concept ARTHRITIS, that doesn’t suffice for you to know which concept is expressed above, and thus you don’t know which proposition is expressed either. It’s unclear which proposition he’s expressing because you don’t know enough about the linguistic community in which this Ghassan travels.4

4 The problem discussed here is distinct from the alleged inconsistency between content externalism and self-knowledge without empirical inquiry, at least as Burge conceives of that allegation. (See, e.g., Burge
The same is true if I introduce a new Ghassan, a Ghassan who isn’t actual and who isn’t in the counterfactual community that refers to tharthritis. Call him “Ghassan2”. I’ll tell you that his community has experts on rheumatoid conditions and these experts refer to some of these conditions by saying “arthritis”. Is that enough information for you to determine which proposition Ghassan2 expresses when he says “Arthritis is common in my family”?

It’s not. You need to know which concept he expresses by “arthritis”; and, according to the tharthritis thought experiment, you don’t know this unless you know that norms and practices in Ghassan2’s community determine that the concept expressed by “arthritis” is thus and so. When it isn’t stipulated as part of the thought experiment that one’s local experts use “arthritis” this way or that, then we have to find out about the norms and practices surrounding the word. If we don’t know who has intellectual authority over the word and who doesn’t, we have to find out and see how the former use “arthritis”. If they use the word differently among themselves, we need to know how diverse practices and norms determine content. And so on.

The point, of course, is that if content externalism is true, then this is where we stand with respect to any (attempted) statement of content externalism. I’ll take claims that include “concept” as the test case.

Recall Burge’s content externalist claim:

[CE] If C is a theoretical concept and “C” expresses C in a community at a time, then the content of C is determined by intellectual norms and practices surrounding C and “C” in the community at the time.

1988) Burge denies that alleged problem by arguing that one may accurately self-attribute beliefs without knowing which propositions those beliefs express. (Ibid 662) But he accepts that empirical inquiry may be necessary to explicate a belief. The problem here is exactly that we (Ghassan included) don’t know which proposition Ghassan expresses without empirical inquiry. I’ll discuss this difference further in section 6.
Imagine one of the Chaunceys uttering CE. Which proposition does he express? As with the Ghassans, we don’t know. If it’s IChauncey, then he’s talking about some subset of Iconcepts (the Iconcepts expressed by theoretical terms); if it’s EChauncey, then he’s talking about a subset of the Econcepts (the Econcepts expressed by theoretical terms). As we saw above, EChauncey’s claim is plausibly true while IChauncey’s claim is plausibly false. But unless we know which Chauncey uttered CE, we don’t know which proposition has been uttered, and we don’t know whether the utterance expresses something true or false. Moreover, if I introduce a new Chauncey, Chauncey2, and you don’t know anything about his community except that they speak a language like English and use the word “concept”, you won’t know what he says when he utters CE either. As with Ghassan2, if you don’t know what the relevant intellectual norms in Chauncey2’s community are, you have to find out, study them, and see how they determine the referent of “concept”; if authoritative uses differ, you have to know how that affects which content is determined; and so on. Until this is all sorted out, we just don’t know what Chauncey2 is saying when he utters CE.

And this is the case if Chauncey2 is actual. In order to find out which proposition he expresses when he utters CE, we should know what our norms with respect to “concept” and CONCEPT are and how they determine the content of CONCEPT. And this is true even if Chauncey2 is himself a content externalist. Until we know which content our norms determine for CONCEPT—and, given the multiform disputes over the nature of concepts, I take it to be plausible that this is something we do not know—, we just don’t know what a content externalist is saying when she utters CE. Or, this is our situation if content externalism is true, anyway.
It’s not as bad as it might be, though. It’s not that we can’t figure out what Burge or Chauncey2 is saying; it’s just that we haven’t yet. We first have to figure out what the norms are, how they determine the referent of “concept”, etc. Perhaps, then, Burge’s claim is no worse off than other claims that feature theoretical concepts, claims like “Ghassan has arthritis”. All we have to do to figure out what’s being expressed by “Ghassan has arthritis” is find out whether the experts in Ghassan’s community think that “arthritis” refers to a rheumatoid ailment that applies to joints only or one that applies to both joints and thighs. The propositions that the sentence might express may be rather neatly constrained, and their differences might not matter much for many of our purposes.

But the point here isn’t that we can’t figure out what the content externalist is saying; it’s just that she hasn’t provided us with enough information to understand her claim. And this is true by her own lights. It’s by her view that her utterance of “concept” has its meaning fixed by some social norms and practices, and it was by appeal to Burge’s thought experiment that we illustrated it.

But these difficulties are indicative of something more troubling. Why don’t we know what the content externalist is saying? Because we don’t know what the crucial term, “concept”, refers to. It may refer to Econcepts, Iconcepts, mental phenomena the contents of which are metaphysically basic (i.e. not determined by anything else), and so on. But suppose there is some fact of the matter about how our concepts have their contents determined. Or, if you don’t like that, suppose there is some fact of the matter about how our concepts are individuated.5 Or,

---

5 If some content externalist would deny that there is some such fact of the matter, she should (at least) explain what she aims to do in proposing a theory of the individuation of concepts. If it turns out that the
alternatively, suppose there is some individuation of contentful mental phenomena that best serves our aims of explaining intelligent human behavior. Call the entities that are so determined or individuated6 “concepts” (rather than “Iconcepts” or “Econcepts” in the terms of this paper). The worry I’m raising here is that we don’t know whether CE expresses a proposition about concepts. Applying externalism to CE, we see that the content externalist may, according to her own theory, be referring to Econcepts when she says “concept”, and it may be that Econcepts aren’t concepts. Just as counterfactual Ghassan’s community refers to tharthritis when they ought to refer to arthritis7, it may be that we refer to Econcepts or Iconcepts when we ought to refer to concepts. If so, then applying externalism to CE has the consequence that CE doesn’t express a proposition about concepts, the mental phenomena we ought to be studying.

Indeed, most of us would bet that our present conception of concepts is more like THARTHITIS than ARTHRITIS—it’s preliminary, rudimentary, and awaiting many further refinements. In fact, as we saw in section 2, Burge seems to think the same. He gives us reason to think that we have, at best, only a tentative grip on the nature of concepts. Concepts on his view depend on norms and practices, but we know very little about these; indeed, he explicitly says we’re in no position to provide a detailed account of the abilities that underwrite possession of some concept. (Burge 2003, 518) But if we’re using a rudimentary, unrefined

---

6 That is, determined or individuated according to the true or most useful theory.

7 Recall that as Burge has it, medical science in the counterfactual community is underdeveloped: “…no one has ever isolated arthritis as a specific disease, or syndrome of diseases.” (1986a, 5)
concept and this is the concept that CE expresses, then CE doesn’t refer to concepts; it’s a claim about something more like tharthritis.

Thus, when we apply content externalism to attempted expressions of content externalism itself, there is no guarantee that the resulting claim expresses a proposition about concepts. Indeed, Burge himself should think it’s unlikely. That is bad. Presumably, content externalism is supposed to be a theory about concepts, not just the referent of “concept” as we happen to use it, no matter how confused we are.

We’re now in a position to summarize and fortify the arguments of this section. First, grant that there’s a fact of the matter about how our concepts have their contents determined—e.g., as internalists say or as externalists say—or about how they should be individuated. Call the entities that have their contents so determined or that are so individuated “concepts”. We may then offer a puzzle as follows.

[P1] If content externalism is true, then our utterances, inscriptions, etc. of CE express a truth about concepts.

[P2] If content externalism is true, then the word “concept” in our utterances, inscriptions, etc. of CE does not refer to concepts.

[P3] If the word “concept” in our utterances, inscriptions, etc. of CE doesn’t refer to concepts, then our utterances, inscriptions, etc. of CE don’t express a truth about concepts.

Suppose content externalism is true. Then by P1 our utterances of CE express a truth about concepts; and by P2, the word “concept” in our utterances of CE does not refer to concepts. But if the word “concept” in our utterances of CE doesn’t refer to concepts, then by P3, our utterances of CE don’t express a truth about concepts. But if it’s not true that our utterances of CE express a truth about concepts, then by P1, content externalism is not true.
Perhaps there are ways to reject P1 or P3, but it shouldn’t be easy. With respect to P1, recall that we’ve referred to CE for concreteness, but our argument should go through for any attempted expression of content externalism. After all, any such expression will have to use words that (one hopes) refer to mental phenomena. I think one could give reasons for rejecting P3; but, I’ll take it that if one would like to reject P3 as a way of avoiding the conclusion, one must show how this goes: the ways in which one might express a truth about X without referring to X wouldn’t seem to offer grounds for recovering externalism. I’ll focus on P2, then. We can argue for it as follows.

1. If our norms reflect an immature, preliminary, to-be-revised understanding of contentful mental phenomena, then our uses of “concept” do not refer to concepts.

Compare that claim with this one.

(1*) If the counterfactual community’s norms reflect an immature, preliminary, to-be-revised understanding of rheumatoid conditions, then their uses of “arthritis” do not refer to arthritis.

In the counterfactual community, rather, they refer to tharthritis, and arthritis ≠ tharthritis. If one accepts the conclusion of the tharthritis thought experiment, one accepts 1*. Since the tharthritis thought experiment is taken to justify externalism, I take it that externalists are indeed committed to 1*. But if the content externalist is committed to 1*, she has no clear grounds for rejecting 1. Prima facie, the reasoning here is no different from that underwriting a crucial externalist argument. If an externalist thinks she can reject 1 while accepting 1*, she must show us how.

---

8 Or, that’s what we should take from the tharthritis thought experiment, anyway. Recall again Burge’s characterization of medical science in the counterfactual community: “…no one has ever isolated arthritis as a specific disease, or syndrome of diseases.” (1986a, 5) This determines that they fail to refer to arthritis. If you think otherwise, then you already disagree with the claim against which this paper argues, and I needn’t convince you.
2. Our norms reflect an immature, preliminary, to-be-revised understanding of contentful mental phenomena.⁹

3. Thus, our uses of “concept” do not refer to concepts.

P2 looks to be established. If P1–P3 are all accepted, then content externalism is true just in case it is not true. If we accept content externalism, we face a paradox.

Of course, if the externalist could simply stipulate that “concept” expresses her understanding of it, then she could express her view as she pleases and insist that it’s about concepts. But prima facie this would be to say that the concept she expresses with “concept” is fully determined by her conception of concepts. That is, this would seem to be a commitment to internalism for CONCEPT.

5. What could CE teach us?

Even if we’re not sure whether CE is about concepts, one might still think it has something important to teach us about minds and mental content. I’ll argue here that if we knew CE to be true, we’d learn only a contingent sociological fact. All we would learn is that norms and practices in our community determine the content of a concept such that it is true of the concept’s referents (if there are any) that they are (or would be) as content externalism says they are. In short: we’d learn something about intellectual norms and expert practices in our community, namely, that they determine the concept we express by “concept” to be

---

⁹ Recall that Burge himself gives us reason to believe this: “…we are not in a position to provide a very rich and detailed account of the nexus of abilities that go into having a concept…” (Burge 2003, 518) Thanks to an anonymous referee from this journal for pointing out to me that this quotation is relevant here.
we wouldn’t be able to infer that we’ve learned anything about concepts. Given the improbability that our concept is the ‘right’ concept, the inference would be invalid.

Compare again the tharthritis case. We may take from it that “arthritis is a rheumatoid ailment that afflicts joints and thighs” can express a true proposition. It doesn’t in our community because here “arthritis” refers to a rheumatoid ailment that afflicts only joints. But in counterfactual Ghassan’s community, the expert practices make it so that “arthritis” refers to tharthritis. There, then, the sentence expresses a true proposition. What makes the difference? The difference in expert practices. In general, “arthritis is a rheumatoid ailment that afflicts joints and thighs” expresses a true proposition just in case it is uttered, inscribed, etc. in a linguistic community where the norms determine that “arthritis” refers to tharthritis. (Assuming we hold the meanings of the other words fixed, of course.) If we learn that the inscription expresses a true proposition in some community, we learn that that community uses “arthritis” to refer to tharthritis. We don’t learn that arthritis is a rheumatoid ailment of joints and thighs, of course: it isn’t.

As with the inscription, so too, mutatis mutandis, for CE. Under what conditions does CE express a true proposition? Just in case it is uttered, inscribed, etc. in a community where intellectual norms and expert practices determine that “concept” expresses ECONCEPT and refers to Econcepts. These aren’t the truth conditions for the proposition that CE expresses: as we saw above, we don’t know which proposition CE expresses. These are the conditions under which CE expresses a true proposition. And I’m not saying that the proposition expressed is

---

10 There may be other possible referents of “concept” that would also make CE true, but for brevity and simplicity, I’ll just say “Econcepts”.
true only contingently. The issue here, rather, is that contingent sociological facts suffice, by the content externalist’s own lights, to make it the case that CE expresses a true proposition. If CE expresses a true proposition, it is because norms and practices have brought it about that “concept” out of our mouths expresses ECONCEPT.

Can we infer from this that concepts are determined in the way content externalism says? We can’t, just as counterfactual Ghassan can’t infer from the truth of his claim about tharthritis that arthritis is a rheumatoid ailment that afflicts both joints and thighs. If we learn that CE expresses a true proposition, then, we’re not in a position to claim any knowledge of concepts; we can claim only that norms and practices in our community determine it that our uses of “concept” express ECONCEPT. And this would surely be a contingent sociological fact. Our norms could have determined it that “concept” expresses different content. We could have expressed ICONCEPT. Perhaps other linguistic communities trying to figure out how humans have mental content do, did, or will express ICONCEPT. If there are such communities, then utterances of CE in those communities express a false proposition. But this isn’t so thanks to the metaphysical nature of concepts; it’s thanks just to contingent sociological facts about those communities, their experts, and their intellectual norms.

6. How should CE be justified?

If the conclusions of the previous sections are right, then how should one go about showing that CE expresses a true proposition? One could simply show that the concept we express by “concept” refers to Econcepts. According to the content externalist, if this is true, then it is so because intellectual norms and expert practices determine it to be so. Presumably,
then, in order to show that CE expresses a true proposition, a content externalist should show that (i) generally, practices and norms determine the content of a concept in ways X, Y, and Z, (ii) our norms and practices concerning the word “concept” and the concept CONCEPT are A, B, and C, and (iii) by ways X, Y, and Z, norms and practices A, B, and C determine that the content expressed by “concept” in our community is ECONCEPT. What I want to emphasize in this section is that ii is an empirical claim. In order to establish it, one must study our linguistic community and identify norms and practices A, B, and C. When taking it that the concept expressed by “concept” is socially determined, then, the content externalist should think that some empirical inquiry is required to establish her claim. And, she should think that she hasn’t established her claim until she’s shown a claim like ii to be true. But, of course, this isn’t how most take it that content externalism was established. Most11 take it that Putnam’s and Burge’s thought experiments established it. But prima facie neither of these thought experiments suffice to show that any claim like ii is true.12 So the content externalist should deny that Putnam’s and Burge’s thought experiments give sufficient reason to believe the view.13

Before arguing for these points more carefully, let me distinguish the problem I’m pushing here from the charge that according to content externalism, empirical inquiry is

---

12 I don’t take myself to be making any controversial positive claim about the sort of inquiry required to establish ii. Rather, I claim just that posing thought experiments alone (without, e.g. systematically recording responses to them) does not suffice to justify a claim like ii. I thus take it that it’s unnecessary here to consider philosophical discussions of semantic intuitions, of experts’ semantic judgments vs. lay semantic judgments, and of experimental philosophy of language.
13 I take these points to be worth making partly because (1) so many believe that the thought experiments did establish content externalism and (2) few think that content externalism requires empirical justification.
required for self-knowledge.\textsuperscript{14} (See, e.g., Burge 1988, Boghossian 1989) In brief, the challenge concerning self-knowledge is that if the contents of one’s thoughts are determined by one’s social or natural world, then in order for Ghassan to figure out what he’s thinking, he has to undertake an empirical inquiry. While this problem and the problem I’m raising here may be related, they are distinct insofar as the necessary conditions for self-knowledge differ from the necessary conditions for knowing which proposition some utterance expresses. Burge says enough to distinguish these two in his response to the problem concerning self-knowledge. (Burge 1988, 659-662) For we can accurately self-attribute a belief without knowing which proposition that belief expresses.

One clearly does not have first-person authority about whether one of one’s thoughts is to be explicated or individuated in such and such a way. Nor is there any apparent reason to assume that, in general, one must be able to explicate one’s thoughts correctly in order to know that one is thinking them.

Thus, I can know that I have arthritis, and know I think I have arthritis, even though I do not have a proper criterion for what arthritis is. (Burge 1988, 662)

If both actual Ghassan and counterfactual Ghassan have a thought they’d express by saying, “I think I have arthritis in my thigh”, each expresses a true proposition even though they express different propositions and even though neither is in a position to distinguish his thought from his counterpart’s. Actual Ghassan says he’s thinking of arthritis; counterfactual Ghassan says he’s thinking of tharthritis. Both speak truly. According to Burge, we simply needn’t know everything about our conceptual contents in order to know that we have them. “Basic self-knowledge”, according to Burge, “is self-referential in a way that insures that the object of reference just is the thought being thought.” (Ibid 659) Presumably, however, if Ghassan doesn’t

\textsuperscript{14} I’m grateful to an anonymous referee from this journal for making it clear that this ought to be done.
know whether he’s thinking about arthritis or tharthritis, he doesn’t know which proposition his thoughts and utterances express. And, of course, in order to know which proposition a belief expresses, one can’t be ignorant of which proposition the belief expresses. Thus, for Burge at least, while the claim that self-knowledge requires empirical inquiry is false, it may nonetheless be true that empirical inquiry is required to figure out which proposition some belief or utterance expresses.

Again, then, our problem here concerns knowing which proposition is at stake. Compare an utterance of “arthritis is a rheumatoid ailment that afflicts joints and thighs”. Imagine that you’ve found yourself transported from actuality to some counterfactual community and you come upon Ghassan2. He was the Ghassan above about whom we know very little. We know his language is very like English and that in his community, “arthritis” refers to some rheumatoid ailment or other. Ghassan2 tells you he has arthritis in his thigh, and, by way of explanation, he adds the utterance above. How are you to evaluate his claim? What can you do to discover whether he speaks truly?

Since you’re a good content externalist, you have to find out whether some claims like i-iii (above) hold in Ghassan2’s world. Of particular importance for us: you have to find out about norms and practices in Ghassan2’s world; you have to discover what norms and practices in that world contingently happen to be like. No matter how much you know about the philosophical literature on content externalism and even if you can recite the Putnam and Burge thought experiments verbatim in your sleep, you won’t know which proposition Ghassan2 is expressing if you don’t know about intellectual norms and expert practices in his world. This much is clear from the original tharthritis thought experiment. If you don’t know whether
experts in your world are talking about arthritis or tharthritis, then you don’t know whether claims like “I have arthritis in my thigh” are silly or possibly true. If you’re a content externalist, then it’s clear what you need from Ghassan2: contingent information about the relevant experts in his community and what their practices are. If you’re then transported to a different possible world where someone else tells you “I have arthritis in my thigh”, you’ll need to start all over again. Why? Because information that establishes a claim like ii in one possible world won’t help you in a different possible world. That is, it’s contingently true and acquired through empirical inquiry.

You are in the same position if you find yourself transported from actuality to a different possible world and someone says, “concepts have their contents determined by intellectual norms in a linguistic community”. Just as above you didn’t know whether Ghassan2 referred to arthritis or tharthritis, in this case you don’t know whether the speaker refers to Iconcepts, Econcepts, or something else. If you’re to find out, you need more information about the intellectual norms and expert practices in the speaker’s community. If the speaker wants to convince you of her claim, she should give you that information.

Of course, this is true even if you haven’t been transported from actuality. If, instead, you just don’t know in actuality about the norms and practices governing “concept”, then you don’t know whether the proposition expressed by a claim like the above is true or not. If someone wants to convince you that the claim expresses a true proposition, she should furnish you with that information. Given that all this applies to CE, it should be part of the content externalist’s case that she provides empirical information about the norms and practices that
determine the referent of “concept”. Insofar as content externalists haven’t done this, the case for content externalism is inadequate.

Have content externalists supplied us with the needed information? It’s implausible in the extreme. Given that the case for content externalism rests on the Putnam and Burge thought experiments, one would have to argue that (A) philosophers alone are the ‘experts’ who determine the content expressed by “concept” and (B) responses to the thought experiments (through discussion, peer review, published papers, etc.) reveal enough about philosophers’ norms and practices with respect to “concept” to discover what they determine to be its referent. For the sake of space, let me just say that each of these is wildly implausible on both factual and methodological grounds. Moreover, even if one were to justify A and B, it wouldn’t show that the Putnam and Burge thought experiments give us grounds to establish that “concept” in our linguistic community expresses ECONCEPT. Rather, it would show that analyzing the responses would make available data that would confirm CE. But of course, someone would still have to carry out the analysis. As far as I know, it hasn’t been done.

Thus, if we’re still interested in the content externalist claim—as a claim that won’t tell us about concepts but only about whatever we refer to by “concept”—we’ve got a long way to go before we know whether it’s true or not.

7. Reply and rejoinder

This paper turns primarily on applying content externalism to attempted expressions of content externalism. From this application, it draws unattractive consequences. An advocate of content externalism may reply with a shrug: the theory shouldn’t be judged by its applicability
to just one example but by its overall success. In that regard, the reply goes, the points raised in this paper should be taken with a grain of salt and with one eye on externalism’s many successes.\textsuperscript{15}

But expressions of content externalism are not just any examples to which content externalism might be applied. They are not on a par with other sentences to which one might apply the view. Rather, they ought to be the expressions that tell us what content externalism is, what it is about, and under what conditions it would be true. But we’ve seen here that if an externalist attempts to tell us what her view is, there’s little reason to think she’ll succeed. She would like to express a proposition about concepts, but content externalism itself tells us to doubt that she does so. Indeed, it may be that we simply can’t presently conceive of the view. To the extent that we can conceive of content externalism, it seems to tell us to doubt that we can think or talk about it. This isn’t a matter, then, of a theory’s simply mishandling a single, contrived example. It’s a case of a theory such that if it is true, it may be unthinkable. Whereas a theory that mishandles a few examples may still be useful to us, I submit that an unthinkable theory cannot be. If the theory is conceivable but simply unknown to us at present, we’re in no position to evaluate its utility.

8. Conclusion

I’ve argued that if content externalism is true, then claims that attempt to express it are such that (i) we don’t yet know which propositions they express, (ii) they aren’t likely to be about concepts, (iii) they can teach us only contingent sociological truths, and (iv) they require

\textsuperscript{15} I’m grateful to an anonymous referee from this journal for raising this objection.
empirical support that hasn’t yet been delivered. Along the way, I’ve claimed that (I) content 
externalism applies to CE, (II) “concept” expresses a theoretical concept, and (III) content 
externalism is supposed to be a theory about concepts. I defended I and II in section 2, but for 
the sake of space, I must leave III as an assumption.

Further, I’ve suggested in each case that these consequences arise thanks to internalist 
presuppositions. If internalism were true, there needn’t be a question about which proposition 
CE expresses or about whether it concerns concepts rather than Iconcepts; and, if internalism 
were true, thought experiments would carry more epistemic weight toward establishing 
Burge’s thesis. But, of course, if internalism is true for CONCEPT, then content externalism is 
simply false.

But is there no hope for the content externalist? Suppose we say that CONCEPT has its 
content determined by its referent independent of social practices. Burge thinks perceptual states 
are individuated independent of social practices (Burge 1986b, 262-3); perhaps a similar claim 
about the individuation of metacognitive states—i.e. thoughts about mental states—can be 
justified. If so, then the objections raised in sections 4 and 5 would fail. If CONCEPT has its 
content determined by concepts ‘in the world’, then we may rest assured that CE succeeds in 
being about concepts. And if we knew CE to be true, we would indeed learn something about 
concepts.

The worry in section 6 would remain largely intact, however. If concepts are ‘in the 
world’ fixing the content of CONCEPT, then theories of concepts call for empirical investigation 
into those concepts in the world. Thought experiments aren’t enough. And much would need to 
be clarified and articulated on such an account, of course—for one thing, metacognitive states
don’t seem to be individuated in the same ways as perceptual states; but, for the sake of space, this sketch will have to do. Still, if a content externalist can maintain that CONCEPT has its content determined independent of social practices, CE would make a theoretically interesting claim about concepts.

Lastly, let me note that there may be a deeper problem here. The backdrop for this paper is the philosophical project in which we ask how our thoughts get their contents and which contents they get. Theories attempting to answer these questions often disagree in the contents they attribute to a given thought, to a given utterance, to a given inscription. And of course, we express these theories in utterances and inscriptions. Prima facie, however, in order to determine whether some such theory is true, we need to know which proposition is expressed by an utterance or inscription put forth by a theory’s advocates. And prima facie, in order to know which proposition is expressed by some such utterance or inscription, we need to know which contents are expressed—we need to know which theory of content determination is true. There are, of course, a number of lemmata suppressed in this presentation of the situation, but the foregoing, more detailed discussion of content externalism dimly suggests a conclusion of this sort. It’s a disheartening state of affairs: we don’t know what the relevant theories say if we don’t know which of them is true; but of course we aren’t in a position to determine which of them is true if we don’t know what they say. I’m not committed to this conclusion, though; I just wanted to acknowledge that it may be the case. For all I’ve said here, the problems raised for externalism might also hold for internalism and every other theory of content.

NOTES


