Definition

Samuel Z. Elgin*

December, 2020

Abstract

This paper presents a puzzle about the logic of real definition. In particular, I demonstrate that five principles concerning definition (that it is coextensional and irreflexive, that it applies to its cases, that it permits expansion and that it is itself defined) are incompatible. I then explore the advantages and disadvantages of each principle—one of which must be rejected to restore consistency.

Introduction

Since the inception of our discipline, the notion of real definition has occupied a central role—and in no field is its significance more manifest than metaphysics. Debates within ethics, epistemology, and beyond can all be framed as searches for definitions. For, when the ethicist provides a theory of the good, this can be reasonably understood as a putative definition of the good, and when the epistemologist provides a theory of knowledge, this can be reasonably understood as a putative definition of knowledge. In metaphysics, too, definition plays this role; we might describe a theory of personhood as a view regarding the definition of *being a person*, and a theory of modality as a putative definition of necessity and possibility. But in metaphysics alone definition plays not only this external role—as something which characterizes theories or accounts under consideration—but also an internal role—as an object worthy of investigation itself. It is the metaphysician who debates the nature of definition, and (perhaps caught in the grip of a rather ambitious mood) provides an account of what definition itself is: a definition of real definition.

This is not to say that the notion of definition has gone unopposed: far from it. There are any number of reasons why philosophers might object to the framing of our field in this manner. Perhaps some believe that our theories are too varied for definition to unify them in any theoretically interesting way; perhaps reality is too coarse-grained for definition to make

*My thanks to Michael Della Rocca, Catherine Elgin, Nathaniel Imel and Will Rattoff for providing comments on earlier drafts of this paper—as well as the attendees of U.C. San Diego's Semantics Babble for providing helpful feedback. I am especially grateful to Cian Dorr, whose correspondence lead directly to the development of this paper.

the distinctions metaphysicians typically take them to make; or perhaps reifying definition builds in a gratuitous and suspect piece of ontology.[1] One objection, the response to which constitutes the subject-matter of this paper, is that talk of 'definition' is *unintelligible*. I do not here hope to assuage all philosophers caught in the grip of the intelligibility concern—as Lewis aptly said, "any competent philosopher who does not understand something will take care not to understand anything else whereby it might be explained."[2] But it is my aim to provide clarity where, now, there is none: to bring structure to the landscape and, in so doing, uncover a puzzle that has thus far remained hidden. For 'definition,' as it is used by the metaphysician, is (among other things) a theoretical term. As such, one way to investigate it systematically is to uncover the theoretical role that it plays.

This strategy involves applying a received view about the introduction of theoretical terms to 'definition.' This view traces back to the Carnap (1958, 1966) discussion of *Ramsey Sentences*, and was given new life by Lewis (1970). The underlying thought is that a theory (whether scientific or philosophical) containing a new predicate is equivalent to its expanded postulate: the claim that there exists a unique F that performs every function that the predicate is taken to perform within that theory (the *Ramsey Sentence*, in contrast to the expanded postulate, lacks the uniqueness claim). The theory as a whole thus amounts to an expansive existence claim. If it is true, then there exists an F which functions as the predicate is taken to function; if it is false, then there does not exist an F which functions as the predicate is taken to function. Debates between adherents and skeptics of that theory can thus be reframed as a disagreement about the existence of such an F. In the present case, the disagreement between the adherents and skeptics of a theory of real definition can be understood as a disagreement about whether an F exists which theoretically functions as definition is postulated to.

In some respects, Lewis's account can be conceived of as the metaphysical converse to *semantic ascent*.[3] Semantic ascent occurs when an ontological disagreement is transformed into a semantic one. One might, for example, semantically ascend by transforming the debate over whether unicorns exist into a debate about whether the predicate 'is a unicorn' has a nonempty extension. But for Lewis, the direction of transformation is reversed. Instead of framing disagreements about ontology as disagreements about semantics, disagreements about a theory—one which contains a novel and suspect theoretical term—are transformed into disagreements about ontology.

Lewis's account arguably needs qualifications and refinements—refinements that, to his credit, he acknowledged. Not all novel theoretical terms are defined by expanded postulates. Sometimes, a term is defined explicitly; a scientist or philosopher may state what a predicate is taken to mean when it is introduced. In these cases, the meaning of the predicate is

[1] Of course, one might reasonably maintain that definition is a worthy object of investigation without endorsing the (admittedly quite implausible) claim the entirety of philosophy amounts to the search for definitions.

[2] See Lewis (1986).

[3] See Quine (1948, 1960).

arguably given by its explicit definition, rather than by a theory's expanded postulate. This allows for cases in which the expanded postulate is false, and yet the novel predicate refers. But this qualification (well taken though it is) does not undermine Lewis's central point: that debates between adherents and skeptics of a theory can generally be understood as disagreements over whether a given theoretical role is satisfied.

This reframing, however, only takes us so far. Even after a dispute is recognized to be a disagreement over existence, it remains unclear how to adjudicate that disagreement. In the sciences, empirical factors often come into play. Extensive experiments and observations have, for example, undermined theories of phlogiston and aether; nothing exists which performs the theoretical role that those theories claimed. But in metaphysics, empirical evidence often seems less relevant. Thought experiments bear on the theory of personal identity—laboratory experiments do not. There is, however, an iron weapon within the skeptic's arsenal; if it can be shown that a theory's expanded postulate is logically inconsistent, then the skeptic has won. The theory is false and—equivalently—there exists no F as adherent claims. At that fatal point, there are two ways to respond. One might abandon the theory wholesale and adopt an alternative in its place, or—more modestly—one might embrace a consistent fragment of the original theory.[4] The task for the adherent, on the second strategy, is to determine which consistent fragment to embrace.

It is my claim that this is the status of 'real definition.' Once an expanded postulate is constructed for the theory of definition, it can be shown to be logically inconsistent. The available responses are either to reject that theory entirely, or else to embrace a consistent fragment of it. The bulk of this paper concerns the identification of that fragment: the arguments for and against the principles in conflict—one of which must be rejected to restore consistency. These principles could be stated with varying degrees of formalism, but I suspect that a quasi-formal gloss is the most easily intelligible (reserving a version I consider to be logically respectable for the end of this paper). They are the following:[5]

[4]This, I take it, was the strategy advanced by Dorr and Hawthorne (2013) regarding Lewis's notion of relative naturalness. In practice, of course, there are more options than these two. Perhaps an adherent will claim that is some ambiguity within the expanded postulate: so that where there appears to be a contradiction, there is in fact none. For the purposes of this paper, I will only consider logically precise theories—ones that do not admit of ambiguity.

[5]Within this paper, I also assume that classical logic holds. I doubt that that assumption is responsible for this conflict; all inferences used to derive the contradiction are admissible on an intuitionist logic, and while a free logic blocks the penultimate inference, I see no independent reason to adopt a free logic in this context. Here, 'Def(F, G)' is intended to be read as 'F is, by definition, G'.

COEXTENTIONALITY: $Def(F,G) \to \forall x(Fx \leftrightarrow Gx)$

IRREFLEXIVITY: $\neg\exists F(Def(F,F))$

CASE CONGRUENCE: $Def(F,G) \to \forall x(Def(Fx,Gx))$

EXPANSION: $(Def(F,G) \land Def(H,I)) \to Def(F,G^{[I/H]})$

DEFINABILITY: $\exists F(Def(Def,F))$

The expanded postulate for this theory of definition, then, results from conjoining these five principles and replacing occurrences of '*Def*' with a variable bound by a (higher order and uniqued) existential quantifier.

Before proceeding to the conflict at hand, it is worth briefly clarifying what these principles mean. *Coextentionality* states that if F is, by definition G, then F and G are coextensive: an object is F just in case it is G. So, for example, if to be a triangle is, by definition, to be a three-angled polygon, then an object is a triangle just in case it is a three-angled polygon. There are no triangles that are not three angled polygons—nor are there three angled polygons that are not triangles. *Irreflexivity* precludes reflexive definitions. It cannot be that to be a person is, by definition, to be a person, or that justice is, by definition, justice. *Case Congruence* claims that definitions apply to their cases. If to be a brother is, by definition, to be a male sibling, then for John to be a brother is, by definition, for John to be a male sibling. And if to be a moral agent is, by definition, to be bound by the categorical imperative, then for Sarah to be a moral agent is, by definition, for Sarah to be bound by the categorical imperative. *Expansion* licenses the substitution of some definitions within the contents of others. If $\{2\}$ is, by definition, the set containing only the number 2 and the number 2 is, by definition, the successor to the number 1 then $\{2\}$ is, by definition, the set containing only the successor to the number 1. And if to be a bachelor is, by definition, to be an unmarried male and to be unmarried is, by definition, to lack a marriage, then to be a bachelor is, by definition, to be a male who lacks a marriage. *Definability*, lastly, states that there exists a definition of real definition—without taking a stand on what the content of that definition is. It asserts that there is some definition of definition or other; definition is not itself a primitive.

I take it that the commitment to these principles is widespread. As we shall see, this commitment is sometimes made explicit; often, it manifests in practice. Moreover, I have no doubt that many would add further criteria to their preferred expanded postulate: criteria reflecting any additional theoretical work that metaphysicians take definition to perform. But it is enough to begin.

The inconsistency between these principles is brought about in the following way:[6]

i)	$Def(Def, D)$	Definability
ii)	$Def(Def(Def, D), D(Def, D))$	*i*, Case Congruence
iii)	$Def(Def(Def, D), D(D, D))$	*i, ii* Definitional Expansion
iv)	$Def(Def, D) \leftrightarrow D(D, D)$	*iii*, Coextensionality
v)	$D(D, D)$	*i, iv* Classical Logic
vi)	$Def(D, D) \leftrightarrow D(D, D)$	*i*, Coextensionality
vii)	$Def(D, D)$	*v*, *vi*, Classical Logic
viii)	$\exists F(Def(F, F))$	*vii*, Classical Logic
ix)	$\bot$	*viii*, Irreflexivity

The expanded postulate for this theory of definition is logically inconsistent, and is therefore false. Those who would continue operate with a notion of definition must articulate which part of the theory they reject—i.e., must identify at least one of the five principles to abandon—and provide a justification for doing so.

Of course, one argumentative technique is apparent; anyone who accepted four of these principles could employ them to derive the negation of the fifth. But that is no help in determining which four to select. What we seek are independent considerations—ones entirely unrelated to this puzzle—that can guide our hand in determining what to do. It is the discussion of these considerations that will occupy the remainder of this paper. For what it's worth, I suspect that many metaphysicians will be loathe to reject the principles I have the least to say about (*coextensionality* and *irreflexivity*); they are starting points in a theory of real definition. However, I ultimately take no stand on how this puzzle ought to be resolved. What I offer are the advantages and disadvantages of each principle. How to weigh these competing considerations is a task I ultimately leave to the reader.

Coextentionality

Coextensionality amounts to the claim that if F is, by definition G, then an object is F just in case it is G. If to be morally right is, by definition, to comply with the categorical

[6]Let 'D' represent the content of the definition of definition—whatever that content might be.

imperative, then an act is morally right just in case it complies with the categorical imperative. And if to believe that p is, by definition, to be disposed to act as if p, then there are neither cases in which an agent believes that p yet is not disposed to act as if p, nor cases in which an agent is disposed to act as if p yet does not believe that p.

There are several reasons to maintain that *coextensionality* is true. Perhaps the most persuasive is an appeal to philosophical practice. Philosophers regularly dismiss putative definitions on the basis of counterexamples. Because there are plausible cases of justified true beliefs that are not cases of knowledge, many deny that knowledge is, by definition, justified true belief.[7] If *coextensionality* were false, the presence of justified true beliefs that are not knowledge would pose no threat to the view that knowledge is justified true belief. To the extent to which other disciplines engage in the search for definitions, the tacit commitment to *coextensionality* appears widespread. On one conception, many chemists search for the definitions of chemical compounds; for, upon uncovering the molecular structure of a given compound, the chemist reveals what the definition of that compound is. If *coextensionality* were false, this practice would be undermined. A chemist may find instances of a compound C that are not instances of molecular structure M without this undermining the claim that C is, by definition, M.

Another route to *coextensionality* passes through identity. Many maintain that if F is, by definition G, then F and G are identical (that is to say, that definitions are a subset of identity claims).[8] I take it that the general thought behind this is the following: whatever definition is, it ought to be reductive. If molecules are defined by their atomic makeup, then those molecules are nothing over and above that atomic makeup. And if normative properties can be defined in purely non-normative terms, then the normative properties can be reduced to non-normative properties. Accounts of definition that fall short of identity arguably fail these reductive ambitions. How could it be that F reduces to G if F remains distinct from G—as something which independently exists? But if definition entails identity, then Leibniz's Law comes into play.[9] That is to say, if F and G are identical, then they bear the same properties; each bears the property *contains object a within its extension* just in case the other does—and similarly so for all other objects. And, for this reason, F and G are coextensive.

It is worth pausing to consider how weak a commitment *coextensionality* is. The metaphysical orthodoxy is that definition is co-*in*tensional. That is, if F is, by definition, G, then F and G have the same extension in every possible world.[10] Cointensionality is strictly

[7] See, canonically, Gettier (1963).

[8] See Correia (2017) for someone who assumes without argument that this is true.

[9] Strictly, a higher-order analogue of Leibniz's Law comes into play. There have recently been numerous discussions of higher order systems that abandon Leibniz's Law; see Caie, Goodman and Lederman (Forthcoming); Bacon and Russell (2019); Bacon (2019).

[10] It was once widely held that this conditional could be strengthened into a biconditional: that is, F is, by definition G, iff F and G have the same extension in every possible world. However, following examples provided by Fine (1994, 1995*a*), many maintain that there are necessary connections between properties that are not definitions. Nevertheless, these examples do not undermine the conditional above.

stronger than *coextensionality* (at least if we assume the T axiom: $\Box P \rightarrow P$), so those who subscribe to the received wisdom must maintain that *coextensionality* is true. There is, however, a small (yet growing) group of philosophers who deny that definition is cointensional.[11] These philosophers often argue that there is no adequate explanation for the link between definition and modality. Perhaps definitions hold contingently; it may be that in the actual world water is, by definition, the chemical compound H_2O, but that in a different possible world water is, by definition, the chemical compound XYZ. I myself find these arguments deeply unpersuasive, but we need take no stand on their merits here.[12] *Coextensionality* makes no assumptions about extensions in other possible worlds; the actual world will do.

Moreover, I note that the general form of *coextensionality* follows from its propositional instance (according to which if p is, by definition, q, then p holds iff q holds) and *case congruence*. To see why this is the case, take an arbitrary F and G such that F is, by definition, G—and an arbitrary object a. *Case congruence* entails that Fa is, by definition, Ga, and the propositional instance of *coextensionality* then entails $Fa \leftrightarrow Ga$. Because the selection of a was arbitrary, F and G are coextensive. Those committed to *case congruence* and the propositional instance of *coextensionality* are thus committed to *coextensionality* in its full generality.[13]

If there are independent reasons to reject *coextensionality*, I am not aware of them.

Irreflexivity

Irreflexivity is the claim that there are no reflexive definitions. Like *coextensionality*, *irreflexivity* is perhaps best defended by an appeal to philosophical practice. Strange as the literature on personal identity undoubtedly is, I know of no one who claims that Socrates is, by definition, Socrates. And while there are those who have argued that knowledge is primitive, I know of no one who has suggested that knowledge is, by definition, knowledge. If reflexive definitions were admissible, these would be glaring possibilities that we, as a philosophical community, have overlooked. Conversely, our collective refusal to regard these possibilities as legitimate reflects our collective commitment to *irreflexivity*.

Some endorse *irreflexivity*, not because they are overly concerned with our practice, but rather because they maintain that definition is itself defined in terms of another irreflexive relation. In various ways, Rosen (2015), Correia (2017), and Horvath (2017) each propose a definition of definition in terms of grounding: an asymmetric relation of metaphysical dependence.[14] A bit roughly, if F is, by definition, G, then the fact that Fa is grounded

[11]See Leech (2018, Forthcoming); Mackie (2020); Romeros (2019). These discussions are framed in terms of the link between essence and modality, but I take it that they could be restated in terms of the link between definition and modality.

[12]See Elgin (Forthcoming*a*).

[13]I discuss the relation between *coextensionality* and *case congruence* in more depth below.

[14]Correia also proposes an account in terms of relative naturalness. What follows is a rough gloss on

in the fact that *Ga*. If to be morally right is, by definition, to maximize utility, then the fact that an act maximizes utility grounds the fact that it is morally right. Because no facts ground themselves, definition is an irreflexive relation. Notably, each author takes the irreflexivity of definition to be not only a feature, but a virtue. That is to say, they take to be a mark in favor of their respective accounts that they preclude reflexive definitions. This might reasonably be taken to indicate that the commitment to *irreflexivity* runs deep.

What reasons are there to reject *irreflexivity*? Some might due to plausible counterexamples. Contemporary cases of reflexive definitions are scarce, but historical examples are comparatively abundant.[15] A particularly notable example is that of *substance* in Spinoza (1677 (1996)).

Spinoza held that everything which exists is either a substance or a mode of that substance. A substance is something which needs nothing else in order to exist. Many of Spinoza's contemporaries (as well as many of his predecessors and successors) held that there is a multiplicity of substances. In contrast, Spinoza held that there is but a single substance: God. Everything else—the entirety of the world we observe around us—are simply modes (or properties) of God. But for the present discussion, the important point isn't Spinoza's defense of monism, but rather his account of substance. He claimed, "By substance, I mean that which is in itself and is conceived through itself: in other words, that of which a conception can be formed independently of any other conception." Strictly speaking, Spinoza's claim is no violation of irreflexivity. He does not assert 'A substance is, by definition, a substance' (indeed, the term 'substance' appears on only one side of the conditional)—but this might be seen as a quibble over details, rather than substance. Spinoza's account, we might think, is as close to an explicit denial of *irreflexivity* as we can reasonably expect to find. A substance is something which is in and conceived in itself.[16] Those who adopt a Spinozistic conception of substance thus have a reason to reject

their views that warrant further refinement. For example, while Rosen takes grounding to be a relation between facts, Correia holds that it is a relation between generics. I direct those interested in the details of these accounts to the original papers.

[15] One possible exception to the (otherwise remarkably widespread) contemporary endorsement of *irreflexivity* is Jenkins (2011). Jenkins explicitly addresses metaphysical dependence rather than definition, but I suspect her arguments could be restated in terms of definition equally well. She argues that metaphysicians are sometimes tempted to claim that Paul's pain depends upon his C-fibers firing, and are, at other times, tempted to claim that Paul's pain *simply is* his C-fibers firing. These views are compatible if Paul's pain depends upon itself.

[16] There is room for dispute over the status of Spinoza's definitions; it is not entirely clear whether they are real—the definitions of things themselves—or merely nominal—the definitions of words. The quote above somewhat suggests a nominal definition of 'substance.' The phrase 'By substance, I mean that which...' indicates that Spinoza was specifying the meaning of a word. After all, words have meanings, while substance need not. What's more, within his preface to Spinoza's *René Descartes Principles of Philosophy*, Lodewijk Meyer claims, "Definitions are nothing but the clearest explanations of the words and terms by which the things to be discussed are designated." This clearly states that the given definitions are nominal and—since Spinoza consented for the preface to appear in his book—can reasonably be taken to reflect his views. If Spinoza's definition of substance is a nominal definition, it need not violate *irreflexivity*, which is a principle regarding real definition. For a discussion of this point, see Lin (2019).

irreflexivity. Because substance is defined reflexively, there are reflexive definitions.

However, I suspect that most contemporary philosophers who reject *irreflexivity* will do so, not because they have particular counterexamples in mind, but rather because they maintain that identity performs the theoretical work often attributed to definition. This may seem particularly appealing given the recent literature on identification—a targeted reading of sentences of the form 'To be F is to be G,' in which the 'is' shares the logical and modal profile of identity.[17] Along these lines, one might think that there are pragmatic reasons to refrain from uttering sentences of the form 'To be F is, by definition, to be F.' Just as it is infelicitous to respond to 'Who is Bob?' with 'Bob is Bob,' so too it is infelicitous to respond to 'What is virtue?' with 'Virtue is virtue.' But in both cases the answers, although entirely unhelpful, remain strictly true. And so, rather than maintaining that there are no reflexive definitions, it might be argued that everything can be defined reflexively.

It is not entirely clear how to make this objection stick. The defender of definition is free to grant that there is a reflexive and symmetric reading of 'To be F is to be G,' but insist that that is not the same reading as intended by their use of 'To be F is, by definition, to be G.' Such a metaphysician may claim that their use of 'definition' refers to the subset of identity claims that are substantive—and it is a requirement on substantiveness that the sentence not be reflexive.[18] On this use of 'definition,' the commitment to *irreflexivity* isn't a pragmatic matter at all, but rather a semantic one. And it is far from clear what prevents the metaphysician from using the term 'definition' in that way.

There are, then, at least two reasons why some might reject *irreflexivity.* They might, firstly, hold particular counterexamples in mind—perhaps a Spinozistic conception of substance. Secondly, they might insist that definitions are reflexive—because identity is reflexive and performs the theoretical work attributed to definition. But I suspect that many

However, while Spinoza devotes no discussion to the status of definition in the *Ethics*, he does within the *Treatise on the Emendation of the Intellect.* There, he provides criteria for definition which indicates that they are real, rather than merely nominal. He claims that if an entity is uncreated, then its definition must not concern external objects, but that if an entity is created, its definition must specify the cause by which it is created. Moreover, the definition of an entity—whether created or uncreated—reveals all of its essential properties. While these are (at least somewhat) plausible criteria for real definition, they are entirely implausible criteria for nominal definitions. There need be no reason why the nominal definition of 'substance' need reveal every essential property of substances. And so, for this reason, some argue that Spinoza's definitions are real. See, again, Lin (2019). A more sophisticated account—according to which Spinoza's definitions are both real and nominal (in particular, that they take us from nominal definitions to real definitions) occurs within Garrett (2003). It appears, then, that there is a reasonable reading of Spinoza's definition of substance on which it is a real definition—and a violation of *irreflexivity.* Those who define substance reflexively in this manner thus have a reason to deny that *irreflexivity* is true.

[17] See, for example, Dorr (2017); Caie, Goodman and Lederman (Forthcoming); Bacon and Russell (2019); Bacon (2019); Fritz (Forthcoming). Some also refer to these sentences as 'just-is' statements—see Rayo (2013); Linnebo (2014) and 'generalized identities'—see Correia and Skiles (2019); Elgin (Forthcoming*b*). As far as I can tell, these are three labels for the same phenomenon.

[18] For a critique of the literature on identification along these lines, see Cameron (2014).

will find neither alternative particularly appealing, so it is worth directing our attention to *case congruence*.

Case Congruence

Case Congruence is the claim that definitions apply to their cases. If to be even is, by definition, to be an integer divisible by two without remainder, then for four to be even is, by definition, for four to be an integer divisible by two without remainder; and if to be a béchamel is, by definition, to be a roux with milk, then for sauce s to be a béchamel is, by definition, for sauce s to be a roux with milk.

As with *coextensionality* and *irreflexivity*, an initial defense of *case congruence* is made by appeal to practice. If it were false, then it ought to admit of counterexample. It may be, for instance, that to be morally right is, by definition, to maximize utility, and for Tim's act to be morally right is, by definition, for Tim's act to comply with the categorical imperative. I am aware of no philosophers who have made claims along these lines—and I take this to indicate that the tacit commitment to *case congruence* is widespread.[19]

In light of the previously-mentioned relation between *case congruence* and *coextensionality* it is worth distinguishing these principles from one another. They are independent; each could be true while the other is false. Let us stick with the previous case: suppose that to be morally right is, by definition, to maximize utility and for Tim's act to be right is, by definition, for Tim's act to comply with the categorical imperative. This, as noted above, is a situation in which *case congruence* fails.[20] Yet so long as Tim's act *both* maximizes utility *and* complies with the categorical imperative, it is no counterexample to *coextensionality*. After all, Tim's act falls within the extension of both *being morally right* and *maximizing utility*, and the propositions *Tim's act is morally right* and *Tim's act complies with the categorical imperative* are both perfectly true. So it may be that *coextensionality* is true while *case congruence* is false.

We can also construct a case in which *case congruence* is true and *coextensionality* is false. Suppose that for Linda to be a sister is, by definition, for Linda to be a female sibling,

[19]This particular example is somewhat tricky. Parfit (2011) argues that Kantianism and consequentialism (as well as contractarianism) are unified in the sense that the best versions of the three views are the same—advocates for each have been climbing the same mountain from different sides. A Parfitian might claim something close to the example above: the definition of the right is given in consequentialist terms and the definition of Tim's act being morally right is given in Kantian terms, because consequentialism and Kantianism are one and the same (at least when understood correctly). But this is not a counterexample to *case congruence*—precisely because the Parfitian identifies the consequentialist view with the Kantian view. A counterexample would be someone who maintains that the two are distinct, rival views—and while the right is defined in consequentialist terms the claim that Tim's act is right is defined in Kantian terms. I know of no one who subscribes to such a claim.

[20]A bit more precisely, we could simply include the negation of 'For Tim's act to be morally right is, by definition, for Tim's act to maximize utility.' An analogous case in which *coextensionality* is true yet *case congruence* is false could be straightforwardly constructed along these lines.

yet the claim that Linda is a sister is true while the claim that Linda is a female sibling is false. Here, *case congruence* is vacuously true (after all, 'Linda is a sister' is 0-ary, and so has no cases to apply to). Yet, here *coextensionality* is false. Linda falls within the extension of *being a sister* yet does not fall within the extension of *being a female sibling*, despite the fact that to be a sister is, by definition, to be a female sibling. And so, *coextensionality* and *case congruence* are independent from one another.

I take it that *case congruence* is an extremely natural principle; one that many metaphysicians assume without argument. But this is not to say that it has gone entirely unchallenged. A recent (and rather sustained) discussion of this type of view occurs in Fine (2016). The subject-matter of Fine's discussion differs slightly from mine. While I discuss principles of definition, Fine discusses principles of identity: what it is that grounds (or metaphysically explains) the fact that various sorts of entities are identical. It may be, for example, that two sets are identical in virtue of having identical members, and it may be that two people are identical in virtue of having continuous conscience experiences. Many maintain that there is a close connection between definition and identity criteria; it is not uncommon for the notion of real definition to be introduced as something which provides identity conditions. On this conception, those concerned with the definition of Socrates investigate what it takes to be identical to Socrates; and those concerned with the definition of knowledge investigates what it takes for a mental state to be identical to knowledge. But regardless of whether this is the appropriate way to characterize the relation between definition and identity, it is natural to expect discussions of one to translate into discussions of the other.

Much of Fine's discussion concerns the priority between generic and specific identity claims. In particular, he maintains that specific claims are metaphysically posterior to (or hold in virtue of) generic claims. The claim '{Hesperus} is identical to {Phosphorus} because Hesperus is identical to Phosphorus' holds because sets are identical in virtue of having identical members. It is not, in contrast, that 'sets are identical in virtue of having identical members' holds (even partially) because of facts regarding {Hesperus} and {Phosphorus}. This view can be straightforwardly translated to a claim about definition. Along these lines, one might maintain that generic definitions are prior to specific definitions. 'Fa is, by definition, Ga' holds because 'F is, by definition, G' holds; it is not the case that 'F is, by definition, G' holds (even partially) because 'Fa is, by definition, Ga' holds.

This point is orthogonal to the debate over *case congruence*, which takes no stand on issues about priority. *Case congruence* states neither that generic claims hold because of specific claims, nor that specific claims because of generic claims. All that it asserts is that generic claims entail specific ones—that one may infer from 'To be a vixen is, by definition, to be a female fox,' that 'For Wanda to be a vixen is, by definition, for Wanda to be a female fox.'

Fine also presents putative counterexamples to (the analog of) *case congruence*. He maintains that some generic identity criteria do not entail all of their specific instances. That is to say, while a claim about generic identity criteria is true, some of its corresponding instances are false. One such example is the following:

(Euclid) $x = a, y = a \Rightarrow x = y$[21]

The claims that the fact that an arbitrary x is identical to a (or, more accurately, takes a as its value), and the fact that an arbitrary y is identical to a (or takes a as its value) collectively ground the fact that x is identical to y. This is a generic claim—as witnessed by the arbitrary objects description, and the free variables x, y within the formal statement. But although this principle appears plausible enough, one of its instances is the following:

$$a = a, a = a \Rightarrow a = a$$

The fact that $a = a$ and the fact that $a = a$ ground the fact that $a = a$. This is false, Fine maintains, because it violates the Non-Circularity of ground, according to which a fact cannot (even partially) ground itself. In order to retain both the generic identity claim and the principle of Non-Circularity, Fine rejects the view that generic claims entail all of their specific instances.

Another example concerns a non-well-founded set theory. Consider a theory which allows for a single non-well-founded set; set ss contains itself (and only itself), and no other set contains itself. The following is a plausible identity criterion for ss.

$$x \in x, \forall y(y \in x)(y = x) \Rightarrow x = ss$$

The fact that an arbitrary set is a member of itself and the fact that all objects which are members of that set are identical to it collectively ground the fact that that set is identical to ss. As before, the reference to arbitrary objects (as well as the occurrence of free variables) indicate that the criterion is generic, rather than specific. Plausible though this principle is, one of its instances is the following:

$$ss \in ss, \forall y(y \in ss)(y = ss) \Rightarrow ss = ss$$

The fact that ss is an element of ss and the fact that every object which is an element of ss is identical to ss collectively ground the fact that ss is identical to ss. But the fact $\forall y(y \in ss)(y = xx)$ is itself plausibly (at least partially) grounded by $ss = ss$—i.e., the fact that all objects within ss are identical to ss is partially grounded in the fact that ss is identical to ss. This is another violation of Non-Circularity. In order to preserve the generic principle and Non-Circularity, the link between the generic and specific identity claims must be severed.

These are the most compelling examples I have yet come across—which is not to say that I find them to be conclusive. For my part, I am unsure of why we ought to accept a principle

[21] Note that the use of '$\Rightarrow$' here refers to grounding, rather than the material conditional, so that '$A \Rightarrow B$' is to be read as 'Fact A grounds fact B.'

of Non-Circularity for non-well-founded set theories. If sets contain themselves, there may be circular dependence relations. Arguably, those who would preserve Non-Circularity principles ought to reject non-well-founded set theories, rather than the connection between generic and specific identity claims.

In any case, to determine whether these examples provide a reason to reject *case congruence*, we must determine whether their definitional analogues are plausible. That is to say, we ought to replace talk of 'identity' with talk of 'definition,' and determine whether the resulting examples carry weight. Take the following:

(Euclid*) $Def(x,\ a),\ Def(y,\ a) \Rightarrow Def(x,\ y)$

This asserts that the fact that (an arbitrary) x is, by definition, a and the fact that (an arbitrary) y is, by definition, a collectively ground the fact that x is, by definition, y. This simply results from replacing occurrences of '$=$' with occurrences of *Def*: from taking the principle to concern definition, rather than identity.

The analogous instance of this principle is the following:

$$Def(a,\ a),\ Def(a,\ a) \Rightarrow Def(a,\ a)$$

This asserts that the fact that a is, by definition a and the fact that a is, by definition, a ground the fact that a is, by definition, a. But immediately, a crucial difference between identity and definition arises: while identity is reflexive, definition is irreflexive.[22] Because all objects are self-identical, a in particular is self-identical; there is a fact that $a = a$. In contrast, if definition is irreflexive then a is not defined in terms of itself; there is no fact that a is, by definition, a. And because there is no such fact, the fact does not ground itself. Quite generally, it is reasonable to maintain that non-existent facts do not ground anything, including themselves. And so, while Fine's original example relies upon the fact that $a = a$ (a fact which no doubt exists), there is no fact that a is, by definition, a to problematize *case congruence*.

I note, moreover, that Euclid* is not nearly as plausible as Euclid. I see no reason to take the fact that x and y are both defined in terms of a to ground the fact that x is, by definition, y. Even before particular instantiations are considered, this principle clearly leads to circular definitions. If two entities are each defined in terms of a, this principle entails that each is defined in terms of the other. And if y takes the same value as x, reflexive definitions immediately result; $Def(x,\ a),\ Def(x,\ a) \rightarrow Def(x,\ x)$. It seems, then, there are ample reasons to reject Euclid*.

Consider the analogue of Fine's second example, concerning a non-well-founded set theory:

$$x \in x, \forall y(y \in x) Def(y,\ x) \Rightarrow Def(x,\ ss)$$

[22]This point (obviously) assumes that *irreflexivity* is true. Those who reject *irreflexivity* could not make this appeal.

This asserts that the fact that $x \in x$ and the fact that $\forall y(y \in x)$*Def(y, x)* collectively ground the fact that *Def(x, ss)*—i.e., the fact that x is an element of itself and the fact that all members of x are, by definition, x collectively ground the fact that x is, by definition, ss.

I see no reason to accept that this is true. As with Euclid*, it immediately generates reflexive definitions. Because ss is an element of itself, this principle entails that *Def(ss, ss)*—ss is, by definition, ss. If *irreflexivity* is true, then there is no such fact. This fact does not stand in circular grounding relations because it doesn't stand in any grounding relations; facts which do not exist neither ground nor are grounded by anything at all.

The upshot is this: there is difficulty in translating Fine's discussion concerning identity to a discussion concerning definition. His examples originally relied upon the fact that object s is self identical. Because all objects are self identical, such a fact assuredly exists. But after translation, the cases concern the fact that object a is, by definition, a. If *irreflexivity* holds, then there is no such fact. For this reason, it is challenging to make these examples stick.

Case Congruence—the view that generic definitions entail their instances—is a natural assumption about definition; arguably, one that many tacitly endorse. Perhaps it is undermined by counterexamples, but—if so—these counterexamples have not yet emerged.

Expansion

Expansion is the claim that, within the contents of a definition, terms may be replaced by their own definitions. For example, if {Socrates} is, by definition, the set containing only Socrates and Socrates is, by definition, the result of *this* sperm and *that* egg, then {Socrates} is, by definition, the set containing only the result of *this* sperm and *that* egg. And if hydrogen is, by definition, the element with a single proton and a proton is, by definition, the particle made of two up quarks and a down quark, then hydrogen is, by definition, the element with a single particle made of two up quarks and a down quark.

Expansion is a restricted substitution principle. It permits substitution within the definiens—or content of definition—but not the definiendum—or object being defined. This restriction matters because an unrestricted principle (i.e., one which allowed for substitution within both the content and object of definition) immediately generates reflexive definitions. Consider the following unrestricted principle:

A is, by definition, B

C is, by definition, D

$\therefore$ (A is, by definition, B)$^{[C/D]}$

That is to say, if A is, by definition B and C is, by definition D, then any replacement of C with D within 'A is, by definition, B' is permissible. Suppose, for example, that the property of *being a vixen* is, by definition, the property of *being a female fox.* This principle can be employed to derive that the property of *being a female fox* is, by definition, the property of *being a female fox*—a reflexive definition. Quite generally, by allowing the same example to witness the first two conditions it is possible to derive reflexive definitions.[23] *Expansion*, as it is stated, does not license substitution within the definiendum, so it does not conflict with *irreflexivity* in this way.

Expansion is the cousin of *transitivity*—the claim that if A is, by definition B and B is, by definition, C, then A is, by definition, C. Strictly, *expansion* is stronger than *transitivity*; *transitivity* can be considered as the limiting case of *expansion* in which the term being substituted for is the definiens in its entirety. But while *expansion* allows us to 'dive into' the content of definiens and replace some terms with others, *transitivity* does not—it applies only to definiens in its entirety. As such, while *expansion* entails *transitivity*, *transitivity* does not entail *expansion.*

The commitment to *transitivity* is widespread—though typically without argument.[24] It is often taken to be a starting-point in a theory of definition; it is considered a mark in favor of a theory if it can be shown to be transitive. This suggests a potential path to *expansion.* While one can consistently hold that *transitivity* is true while *expansion* is false, it is not at all clear why we should expect *transitivity* to succeed and *expansion* to fail.[25]

Explicit commitment to *expansion* is less common than the commitment to *transitivity* (though its explicit denial is, as far as I know, nonexistent). An exception to this general rule is the following:

> "It should be possible to prove a principle that licenses arbitrary definitional expansion:
>
> $$Def(F,\Phi) \text{ and } Def(G,\Psi) \text{ then } Def(F,\Phi^{\Psi/G})$$
>
> Where $\Phi^{\Psi/G}$ is the result of substituting Ψ for G in Φ...Any account of real definition should license the substitution of definiens for definiendum in a ground to yield a further ground" (Rosen, 2015, pg. 201).[26]

[23]Note that this is slightly weaker than the claim that definition is *reflexive*—i.e., from the claim that every instance of 'A is, by definition, A' is true. There may well be some irreflexive cases; what this principle entails is that reflexivity arises for every term that serves as the content of a definition.

[24]See, most explicitly, Correia (2017); Rosen (2015); Horvath (2017). For endorsements of the transitivity of related phenomena such as ground, essence, and relative fundamentality, see, for example, Fine (1995*b*, 2012); deRosset (2013, 2017); Dasgupta (2016); Berker (2017); Dixon (2018). The closest thing to an explicit disavowal of transitivity occurs in Schaffer (2012). For a reply, see Litland (2013).

[25]In addition, much theoretical work attributed to *transitivity* can only be adequately accomplished by *expansion.* For example, one type of ontological dependence can be understood in terms of definitional containment—see Fine (1995*a*). Entity e ontologically depends upon entity e' just in case e' figures within the definition of e. *Expansion* can be used to derive the transitivity of ontological dependence, but *transitivity* is strictly compatible with the claim that ontological dependence is intransitive.

[26]For another commitment to this type of principle, see Correia and Skiles (2019).

Notably, Rosen claims not only that a substitution principle holds, but that it ought to be *provable* that it holds. This suggests a path toward expansion; we ought to believe it because of its proof. To the best of my knowledge, however, *expansion* does not follow from any of the widely accepted principles about essence or definition. Minimally, I have been unable to prove it from them. Those seeking a proof of *expansion* should look—not to the logic of essence and definition—but to the logic of identity.

Within the recent (and rapidly expanding) literature on higher-order identity, Caie, Goodman and Lederman (Forthcoming) provide a proof of Leibniz's Law. The aim of this proof is not to vindicate Leibniz's Law, but rather to systematically investigate which principles must be abandoned in languages with opaque predicates. As it turns out, this derivation can be modified to prove *expansion*.

Let us adopt a typed, higher-order language with λ-abstraction. Within this language, there are two basic types e, t for the type of entities and sentences respectively, and for any types τ_1, and $\tau_2 \neq e, (\tau_1 \rightarrow \tau_2)$ is a type; nothing else is a type. Monadic first-order predicates can be identified as terms of type $(e \rightarrow t)$, diadic first-order predicates are terms of type $(e \rightarrow (e \rightarrow t))$, etc.. Monadic second-order predicates are of type $((e \rightarrow t) \rightarrow t)$, and monadic third-order predicates are of type $(((e \rightarrow t) \rightarrow t) \rightarrow t)$. The negation operator $\neg$ is of type $(t \rightarrow t)$, and the binary operators $\wedge, \vee, \rightarrow, \leftrightarrow$ are all of type $(t \rightarrow (t \rightarrow t))$. Additionally, this language is equipped with infinitely many variables of every type, as well as the quantifiers $\exists, \forall$ (for every type).

In first-order languages, these quantifiers perform dual functions. They serve both to express generality and to bind the variables occurring within their scope. But in higher-order languages, these tasks are divided: the task of expressing generality is performed by quantifiers and the task of variable binding is performed solely by the λ-terms. Thus, 'there exists an F' is expressed as '$\exists\lambda x.(Fx)$,' rather than '$\exists x(Fx)$.'

Lastly, for each type τ there exists a predicate *Def* of type $(\tau \rightarrow (\tau \rightarrow t))$ which is used to express definitions. The intended interpretation of $\ulcorner Def^{(\tau \rightarrow (\tau \rightarrow t))}(A^\tau, B^\tau)\urcorner$ is 'A is, by definition, B.'

The principles which generate *expansion* (which are are either to be read as schemata with applications in every type, or whose type is contextually evident) are the following:

MATERIAL ABSTRACTION	$Def(\phi, \psi) \rightarrow Def(\lambda x.\phi[x/a], \lambda x.\psi[x/a])$
APPLICATION CONGRUENCE	$Def(F, G) \wedge Def(a, b) \rightarrow Def(Fa, Gb)$
BETA-ETA EQUIVALENCE	ϕ may be replaced by ψ provided ϕ and ψ are $\beta\eta$ equivalent

The derivation of *expansion* proceeds as follows:

i)	$Def(a, b)$	Supposition
ii)	$Def(\phi, \psi)$	Supposition
iii)	$Def(\lambda x.\phi[x/a], \lambda x.\psi[x/a])$	ii, Material Abstraction
iv)	$Def(\lambda x.\phi[x/a](a), \lambda x.\psi[x/a](b))$	iii, Application Congruence
v)	$Def(\phi, \psi[b/a])$	iv, Beta-Eta Equivalence

Therefore, if *material abstraction, application congruence* and *beta-eta equivalence* are all true, then *expansion* is true as well. Those who would reject *expansion* must also reject (at least) one of these three principles.

Beta-eta equivalence itself follows from two principles: the claim that identification is preserved under $\beta\eta$ conversion and Leibniz's Law. For example, preservation under $\beta\eta$ conversion entails that $\lambda x.Fx(a) = Fa$ and, because these propositions are identical, Leibniz's Law then entails that one term can be substituted for the other. Those who reject *expansion* by rejecting *beta-eta equivalence* must also reject either the claim that identification is preserved under $\beta\eta$ conversion or Leibniz's Law.

The claim that identification is preserved under $\beta\eta$ conversion is an orthodox principle of higher-order logic. Its most sustained defense occurs in Dorr (2017), and I have little to add to that defense. The strongest argument *against* this principle arises from the view that propositions are structured—a view radially incompatible with the claim that identity is preserved under $\beta\eta$ conversion. A central commitment of structured propositions is that $Fa = Gb \rightarrow F = G$; if the proposition that Fa is identical to the proposition that Gb then F is identical to G. $\beta\eta$ conversion entails $\lambda x.Rxx(a) = \lambda x.Rxa(a)$: the proposition that object a stands in relation R to itself is the same as the proposition that object a stands in relation R to a. On the structure proposition view, this entails that $\lambda x.Rxx = \lambda x.Rxa$: the property of standing in relation R to oneself is identical to the property of standing in relation R to a. This is obviously absurd—the two properties need not even be coextensive. Those who endorse structured propositions, then, have a reason to reject the claim that identity is preserved under $\beta\eta$ conversion—and, correspondingly, a path to rejecting both Beta-Eta Equivalence and *expansion*.

However, there is an independent reason to reject structured propositions: the Russell-Myhill problem. Another (and more general) commitment of structured propositions is that syntactic differences correspond to propositional differences. The fact that 'Sarah is to the left of John' differs syntactically from 'John is to the right of Sarah' entails that the two sentences correspond to different propositions. The problem, roughly, is that for every collection of propositions it is possible to construct a sentence asserting that precisely those propositions are true. For this reason, there is a mapping from the powerset of propositions to a unique sentence (i.e., there is a mapping from each collection of propositions to the sentence asserting that precisely the elements of that collection are true). If each sentence itself corresponded to a unique proposition, then there would be a mapping from each element of the powerset of propositions to a unique proposition. But Cantor's Theorem states that there is no such mapping. For any set s there is no mapping from every element of the powerset of s to a unique element of s. And so, it cannot be that every syntactic difference corresponds to a propositional difference.

Moreover, Fritz (Forthcoming) has recently demonstrated that there are other entities—ones with the resources to evade the Russell-Myhill problem—which can perform much of the theoretical work often attributed to structured propositions.[27] A bit roughly, instead of appealing to the structured proposition that Fa, we may appeal to the bihaecceity $R = \lambda X^{e \to t}.\lambda x^{e}.(X = F \land x = a)$: a relation that property F stands in to object a and that no property stands in to any other object. These entities are proxies for structured propositions, but are not themselves propositions; R is a relation between properties and objects, and is therefore not truth-evaluable. As it turns out, these proxies resolve the Russell-Myhill, and are compatible with the claim that identity is preserved under $\beta\eta$ conversion.[28]

The upshot is this: those who endorse structured propositions ought to reject the claim that identity is preserved under $\beta\eta$ conversion. This might seem appealing, as it facilitates the rejection of both Beta-Eta Equivalence and *expansion*. However, the structured-proposition view is deeply flawed (and may not even be consistent)—and at least some of the work often attributed to it can be performed by other sorts of structures. Ultimately, then, I doubt that the present dilemma ought to be resolved in this way.

Leibniz's law also holds substantial appeal. If Hesperus is identical to Phosphorus, then for Hesperus to be a planet is for Phosphorus to be a planet—and if Cicero is identical to Tully, then for Cicero to be an orator is for Tully to be an orator. The strongest reason to reject Leibniz's Law arises from opaque predicates.[29] It may be that early Babylonians believe that Hesperus appeared in the evening sky while denying that Phosphorus appeared in the evening sky. If this is so, then 'Hesperus' may not be replaceable by 'Phosphorus' in some contexts—despite the fact the two are identical.[30] Of course, many continue to endorse Leibniz's law; it is an extremely natural principle—difficult cases notwithstanding.

Application congruence allows for the combination of two definitions into one. If to be human is, by definition, to be a rational animal and Aristotle is, by definition, the result of *this* sperm and *that* egg, then for Aristotle to be human is, by definition, for the result of

[27]In particular, Fritz demonstrates their use in resolving puzzles of ground.

[28]I direct those interested in *how* this resolves the Russell-Myhill to Fritz's original paper. Roughly, we need only assume that the cardinality of proxies is higher than the cardinality of their relata: in the present case, that the cardinality of relations between properties and objects is higher than the cardinality of properties and of objects.

[29]Indeed, opacity might even be defined as any violation of Leibniz's Law.

[30]I note, however, that there are ways of resolving Frege's puzzle that do not involve rejecting Leibniz's Law. See, for example, Dorr (2014).

this sperm and *that* egg to be a rational animal. Application Congruence strongly resembles *case congruence*, and many reasons to accept (and reject) *application congruence* as well. For the moment, suffice it to say that I can think of no plausible instances in which it fails.

Material abstraction is the near converse of *case congruence*. Just as *case congruence* allows one to infer that 'For Linda to be a sister is, by definition, for Linda to be a female sibling' from 'To be a sister is, by definition, to be a female sibling,' so too *material abstraction* allows one to infer 'To be a sister is, by definition, to be a female sibling' from 'For Linda to be a sister is, by definition, for Linda to be a female sibling.' The underlying thought is that when a term appears in both the definiendum and definiens—within both the object and content of analysis—then that term is not responsible for the definition in question. That is to say, there is a plausible a *non-circularity* criterion on definition. While terms can (and do) appear in both the object and contents of definitions, they cannot appear *essentially* in both the object and content—they are not the reason a given expression constitutes a definition.[31] And, because these terms are inessential, they can be abstracted away.

To my mind, the defense of both *application congruence* and *material abstraction* are defeasible. They are not knock-down considerations. One path to the resolution of the problem at issue is the rejection of *expansion*. However, this rejection must be accompanied by the rejection of *application congruence* or *material abstraction* (or *beta-eta equivalence*), as these principles entail that *expansion* is true.

Definability

Definability is the claim that there exists a definition of definition. Definition does not rank among the primitive relations—it is defined in terms of something or other. I suspect that (at least to some) this principle seems relatively controversial. On one interpretation, definition forms a bedrock of our discipline: a foundation upon which other philosophical accounts rest. And so the contention that definition is itself primitive is not entirely implausible. Moreover, while the previous principles appeared to be implicit in philosophical practice, this is not so for *definability*. There is no reason to suspect that practicing ethicists, epistemologists and the like tacitly assume that real definition is itself defined.

Nevertheless, numerous philosophers maintain that *definability* is true. Typically, this occurs because philosophers provide an account of definition.[32] That is to say, philosophers defend a particular view about what the definition of definition is, and are thereby committed to the claim that definition has some definition or other. Correspondingly, one defense of *definability* is parasitic on any argument that they provide. A reason to support their views constitutes a reason to endorse *definability*.

[31] At least, they can appear in both the object and content of definition on the assumption that *case congruence* is true.

[32] As before, see Rosen (2015); Correia (2017); Horvath (2017). I direct those interested in the content of these views to their original papers.

I believe that there is a further reason to support *definability*—but one that is (or at least ought to be) controversial.

Definitions provide answers to metaphysical-why questions. Several things might be intended by a question like 'Why is Fred a bachelor?' Often, it might be used to enquire into the reason for Fred's marital status. In these cases, responses like 'Because he has not yet fallen in love' are appropriate. But there is a metaphysical reading of this question as well—one concerning what it is in virtue of that Fred counts as a bachelor–and it is here that an appeal to definition is made. To the metaphysician, the response 'Because Fred is an unmarried male and to be a bachelor is, by definition, to be an unmarried male' seems as satisfying an answer as any.

In a similar manner, the definition of definition provides an answer to metaphysical-why questions. Let us suppose, for the sake of a concrete example, that to be morally right is, by definition, to maximize utility. It seems a reasonable question to enquire *why* the right is defined as it is: what makes it the case that the right is defined in terms of that which maximized utility rather than that which cultivates the virtues. And just as the answer to 'Why is Fred a bachelor?' naturally appeals to the definition of *being a bachelor*, so too the answer to 'Why is the right, by definition, maximizing utility?' naturally appeals to the definition of definition. The reason the right is defined in terms of maximizing utility is that it stands in the appropriate relation to maximizing utility: a relation articulated by the definition of definition.

Those who accept *definability* have the resources to metaphysically explain why it is properties and relations are defined as they are; they can appeal to the definition of definition to provide such an account. In contrast, those who reject *definability* cannot respond in this way. And so, one reason to accept *definability*—beyond the appeal of particular accounts of definition—is that it provides resources for metaphysical explanations that we seek.

There is, however, a reason to reject *definability*: one so initially compelling that it suggests that the preceding discussion ought to have been curtailed. As stated, *definability* claims that a relation—in particular, the relation of definition—stands within its own extension (while remaining agnostic as to what it stands in that relation to). But there is a strong reason to deny that any property or relation falls within their own extension: the Russell Paradox. For, if properties are contained within their extension, it is natural to maintain that there is a property of *being a property that is not contained within its own extension*: a property that falls within its own extension just in case it does not. This problem can be avoided by denying that properties are the types of things that can fall within their own extension. And if *no* property or relation falls within its own extension, then definition does not fall within the extension of definition, and so we ought to reject *definability*.

This, as I said, is an extraordinarily compelling point. It is also false. Of course, there are numerous ways we might attempt to obviate the Russell Paradox, but the *obvious* method leaves the present puzzle intact. While outright contradiction is avoided, the conflict between the five principles at issue remains. What is this obvious method? To adopt a typed higher-order language in which the claim that properties fall within their own extension are strictly ungrammatical, and so inapt for truth or falsity.

Fortunately, we have already encountered such a language, so no new formalism is required. As before, let us assume that there are two basic types, e and t (for the types of entities and sentences respectively), and that for any types α, β, $(\alpha \to \beta)$ is a type, and nothing else is a type. We allow for infinitely many variables of every type, and the corresponding λ abstracts needed to bind them. Furthermore, for any type α there exists a predicate *Def* of type $\alpha \to (\alpha \to t)$ with the intended interpretation that $Def^{\alpha \to (\alpha \to t)}(A^{\alpha}, B^{\alpha})$ asserts that A is, by definition, B. Because this language is typed, the Russell Paradox is avoided. The only additional symbolism—which I introduce solely to reduce the length of types in the principles and subsequent derivation—is α^2 (for a generic type α) which is shorthand for $\alpha \to (\alpha \to t)$. Relatedly, $(\alpha \to t)^2$ is shorthand for $(\alpha \to t) \to ((\alpha \to t) \to t))$.

With this language in place, the five principles at issue can be stated in a logically precise manner. Strictly, these principles become schemata with applications for each type α. In cases where the type is not explicitly mentioned, it is contextually evident.

COEXTENSIONALITY: $Def^{t^2}(P^t, Q^t) \to (P^t \leftrightarrow Q^t)$

IRREFLEXIVITY: $\neg\exists\lambda X^{\alpha}.Def^{\alpha^2}(X, X)$

CASE CONGRUENCE: $Def^{(\alpha \to t)^2}(F^{\alpha \to t}, G^{\alpha \to t}) \to Def^{t^2}(F^{\alpha \to t}(a^{\alpha}), G^{\alpha \to t}(a^{\alpha}))$

EXPANSION: $(Def^{\alpha^2}(F^{\alpha}, G^{\alpha}) \wedge Def^{\beta^2}(H^{\beta}, I^{\beta})) \to Def^{\alpha^2}(F^{\alpha}, G^{\alpha\ [\mathrm{I/H}]})$

DEFINABILITY: $\exists\lambda X^{\alpha^2}.Def^{\alpha^2 \to (\alpha^2 \to t)}(Def^{\alpha^2}, X)$

Most of these amount to the reframing of the original principles in a paradox-free language. *Definability*, for example, amounts to the claim that there is a higher-order definition for each lower-order definition. *Coextensionality*, however, has been restricted to its relevant application: it is stated for the case of terms of type t (the only type relevant to the derivation of this puzzle) rather than for all types generally.

The framing of this puzzle within a typed language offers another potential resource. It may be that different principles are rejected for different types. Perhaps, for example, *Definability* is to be rejected for the predicate $Def^{e \to (e \to t)}$ while *expansion* is to be rejected for the predicate $Def^{(e \to t) \to ((e \to t) \to t)}$. However, I can think of no reason to reject different principles for different types, so I merely note that it is an option in logical space.

Within this language, the conflict can then be derived in the following way:

i)	$Def^{\alpha^2 \to (\alpha^2 \to t)}(Def^{\alpha^2}, D^{\alpha^2})$	Definability
ii)	$Def^{(\alpha^2 \to (\alpha^2 \to t)) \to ((\alpha^2 \to (\alpha^2 \to t)) \to t)}(Def^{\alpha^2 \to (\alpha^2 \to t)}, D^{\alpha^2 \to (\alpha^2 \to t)})$	Instance of i
iii)	$Def^{t^2}(Def^{\alpha^2 \to (\alpha^2 \to t)}(Def^{\alpha^2}, D^{\alpha^2}), D^{\alpha^2 \to (\alpha^2 \to t)}(Def^{\alpha^2}, D^{\alpha^2}))$	Case Congruence, ii
iv)	$Def^{t^2}(Def^{\alpha^2 \to (\alpha^2 \to t)}(Def^{\alpha^2}, D^{\alpha^2}), D^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2}))$	Expansion, iii
v)	$Def^{\alpha^2 \to (\alpha^2 \to t)}(Def^{\alpha^2}, D^{\alpha^2}) \leftrightarrow D^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2})$	Coextensionality, iv
vi)	$D^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2})$	Classical Logic i, v
vii)	$Def^{t^2}(Def^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2}), D^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2}))$	Case Congruence, i
viii)	$Def^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2}) \leftrightarrow D^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2})$	Coextensionality, vii
ix)	$Def^{\alpha^2 \to (\alpha^2 \to t)}(D^{\alpha^2}, D^{\alpha^2})$	Classical Logic, $vi, viii$
x)	$\exists \lambda X^{\alpha^2}.Def^{\alpha^2 \to (\alpha^2 \to t)}(X, X)$	Classical Logic, ix
xi)	$\perp$	Irreflexivity, x

The upshot, then is this: there was a presumptive concern regarding *definability*. It appeared to assert that a relation—in particular the relation of real definition—fell within its own extension. This naturally gives rise to paradox, and provides an initial reason to reject *definability*. However, once we shift into a typed language, the threat of paradox is removed, and yet the present conflict remains. And so, if there is a reason to reject *definability*, it is not due to the threat of paradox.

Conclusion

I close by returning to where we began: a discussion of the expanded postulate for a theory of definition. I have no doubt that some readers suspect that this expanded postulate (whatever it may be) constitutes the definition of real definition. What definition *is* is that relation that performs the theoretical work attributed to real definition. And so, once we identify what that work consists of, we will thereby have identified what the definition of definition is.

Lewis (1970)'s original work suggests that this is incorrect. Note that his account concerns how to define theoretical *terms*, rather than properties and relations. That suggests that it provides nominal definitions, rather than real definitions. The expanded postulate constitutes a nominal definition of 'Definition' rather than a real definition of definition. It merely specifies what the word means as used by the metaphysician.

My own view is that matters are not so straightforward. What an expanded postulate is is the formal description of the theoretical function that a property (or relation) performs. Properties or relations which are defined in terms of their expanded postulate are thus those which are functionally defined. If definition is one such property—if it is functionally defined—then its expanded postulate provides its real definition. In contrast, if definition is not functionally defined, then its expanded postulate provides a merely nominal definition. The debate over the relation between definition and its expanded postulate can thus be understood as a debate over whether definition is itself functionally defined.

There is a conflict between the principles *coextensionality, irreflexivity, case congruence, expansion* and *definability.* Each holds at least some measure of initial appeal, and while there are modest reasons to reject some, many have a great measure of support.

Perhaps some will respond to this result by abandoning the theory of definition wholesale. I myself am skeptical of that approach—the difficulty of rejecting one principle is no reason to reject five. Although I ultimately take no stand on how this puzzle ought to be resolved, something must be done; the contradiction cannot be allowed to stand.

References

Bacon, Andrew. 2019. "Substitution Structures." *Journal of Philosophical Logic* 48:1017–75.

Bacon, Andrew and Jeffrey Russell. 2019. "The Logic of Opacity." *Philosophy and Phenomenological Research* 99(1):81–114.

Berker, Selim. 2017. "The Unity of Grounding." *Mind* 127(507):729–77.

Caie, Michael, Jeremy Goodman and Harvey Lederman. Forthcoming. "Classical Opacity." *Philosophy and Phenomenological Research* .

Cameron, Ross. 2014. "On the Lack of Direction in Rayo's *The Construction of Logical Space*." *Inquiry* 57(4):427–41.

Carnap, Rudolph. 1958. "Observational and Theoretical Language." *Dialectica* 12(3-4):236–48.

Carnap, Rudolph. 1966. *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*. Basic Books.

Correia, Fabrice. 2017. "Real Definitions." *Philosophical Issues* 27(1):52–73.

Correia, Fabrice and Alexander Skiles. 2019. "Grounding, Essence and Identity." *Philosophy and Phenomenological Research* 3:642–70.

Dasgupta, Shamik. 2016. "Metaphysical Rationalism." *Noûs* 50(2):379–418.

deRosset, Louis. 2013. "Grounding Explanations." *Philosopher's Imprint* 13(7):1–26.

deRosset, Louis. 2017. "Grounding the Unreal." *Philosophy and Phenomenological Research* 95(3):535–63.

Dixon, Scott. 2018. "Upward Grounding." *Philosophy and Phenomenological Research* 97(1):48–78.

Dorr, Cian. 2014. Transparency and the Context-Sensitivity of Attitude Reports. In *Empty Representations: Reference and Non-Existence*, ed. Manuel Garcia-Carpintero and Genoveva Marti. Oxford University Press pp. 25–66.

Dorr, Cian. 2017. "To be F is to be G." *Philosophical Perspectives* 30(1):39–134.

Dorr, Cian and Johnathan Hawthorne. 2013. Naturalness. In *Oxford Studies in Metaphysics*, ed. Karen Bennett.

Elgin, Samuel. Forthcoming*a*. "Essence, Modality and Identity.".

Elgin, Samuel. Forthcoming*b*. "The Semantic Foundations of Philosophical Analysis.".

Fine, Kit. 1994. "Essence and Modality." *Philosophical Perspectives* 8:1–16.

Fine, Kit. 1995*a*. "Ontological Dependence." *Proceedings of the Aristotelian Society* 95:269–90.

Fine, Kit. 1995*b*. "Ontological Dependence." *Proceedings of the Aristotelian Society* 95(1):269–90.

Fine, Kit. 2012. A Guide to Ground. In *Metaphysical Grounding*, ed. Fabrice Correia and Benjamin Schnieder. Cambridge University Press pp. 37–80.

Fine, Kit. 2016. "Identity Criteria and Ground." *Philosophical Studies* 173(1):1–19.

Fritz, Peter. Forthcoming. "Structure by Proxy with an Application to Grounding." *Synthese* .

Garrett, Aaron. 2003. *Meaning in Spinoza's Method*. Cambridge University Press.

Gettier, Edmund. 1963. "Is Knowledge Justified True Belief?" *Analysis* 6:121–3.

Horvath, Joachim. 2017. "Philosophical Analysis: The Concept Grounding View." *Philosophy and Phenomenological Research* 97(3):724–50.

Jenkins, Carrie. 2011. "Is Metaphysical Dependence Irreflexive?" *The Monist* 94(2):267–76.

Leech, Jessica. 2018. "Essence and Mere Necessity." *Royal Institute of Philosophy Supplement Metaphysics* 82:309–32.

Leech, Jessica. Forthcoming. "From Essence to Necessity via Identity." *Mind* .

Lewis, David. 1970. "How to Define Theoretical Terms." *Journal of Philosophy* 67(13):427–46.

Lewis, David. 1986. *On the Plurality of Worlds*. Oxford University Press.

Lin, Martin. 2019. *Being and Reason: An Essay in Spinoza's Metaphysics*. Oxford University Press.

Linnebo, Øystein. 2014. "'Just-Is' statements as generalized identities." *Inquiry* 57(4):466–82.

Litland, Jon. 2013. "On Some Counterexamples to the Transitivity of Ground." *Essays in Philosophy* 14(1):19–32.

Mackie, Penelope. 2020. Can Metaphysical Modality Be Based on Essence? In *Metaphysics, Meaning, and Modality: Themes from Kit Fine*, ed. Mircea Dumitru. Oxford University Press.

Parfit, Derek. 2011. *On What Matters*. Oxford University Press.

Quine, W. V. O. 1948. "On What There Is." *The Review of Metaphysics* 2(5):21–38.

Quine, W. V. O. 1960. *Word and Object*. MIT Press.

Rayo, Augustín. 2013. *The Construction of Logical Space*. Oxford University Press.

Romeros, Carlos. 2019. "Modality Is Not Explainable by Essence." 69:121–41.

Rosen, Gideon. 2015. "Real Definition." *Analytic Philosophy* 56(3):189–205.

Schaffer, Jonathan. 2012. Grounding, Transitivity and Contrastivity. In *Metaphysical Grounding: Understanding the Structure of Reality*, ed. Fabrice Correia and Benjamin Schnieder. Cambridge University Press pp. 122–38.

Spinoza, Benedictus. 1677 (1996)). *The Ethics*. Penguin Classics.