MODALITY IS NOT EXPLAINABLE BY ESSENCE

BY CARLOS ROMERO

Some metaphysicians believe that metaphysical modality is explainable by the essences of objects. In Section II, I spell out the definitional view of essence, and in Section III, a working notion of metaphysical explanation. Then, in Section IV, I consider and reject five natural ways to explain necessity by essence: in terms of the principle that essential properties can’t change, in terms of the supposed obviousness of the necessity of essential truth, in terms of the logical necessity of definitions, in terms of Fine’s logic of essence, and in terms of the theory of real definitions. I will conclude that the present evidence favours rejecting the hypothesis that modality is explainable by essence.

Keywords: necessity, essentialism, explanation, grounding, real definitions.

I. INTRODUCTION

Fine (1994) has rejected the reduction of essence to \textit{de re} necessity.\footnote{See Torza (2015) for a generalization of Fine’s critique.} But he has also suggested a reduction of modality to essence. Relatedly, Lowe, Hale and others have argued that necessity is grounded in essence. I take these views as cases of the general view that modality is \textit{explainable by} the essences. Here, I examine and reject five of the most natural ways to ground this general doctrine.\footnote{For other doubts about essentialist explanations of modality, see Cameron (2008): 272; Teitel (2018); Wildman (2018); Whittle (2010); Hale (1996) (see Shalkowski 1997 for discussion of Hale’s objection). Hale himself endorses an essentialist explanation of modality: Hale 2002, 2013.}

Before proceeding, let me note that it is not obvious that there \textit{are} essential truths. There are doubts about the notion of \textit{objective definition} (Robertson and Atkins 2018: Section 2), about objectually defining \textit{concrete} objects (Almog 2010), and about what is perhaps the most intuitive case of individual essence—origins (Mackie 2006: ch. 3). Some have tried to explain away the intuitive examples of essentialist truth in terms of semantic conventions (Sidelle 1989) or of cognitive biases (Leslie 2013). Others have coupled the view of essences as cognitive
devices with the rejection of the individuals they are essences of (Ladyman and Ross 2007: 241). Lewis’ contextualist theory also deflates essentialist truth (1986: 252; but see Paul (2004) for an essentialist alternative). Here, however, I ignore these issues: I mean to show that, even if there are essential truths, they do not suffice to explain modality.

II. DEFINITIONAL ESSENTIALISM

Definitional essentialism is the view that construes essence in the model of real definition, instead of as a special case of de re modality. Definitional essentialism has been championed by Fine, Lowe and Hale.

Fine argued that the notion of essence (or nature) should be taken as primitive, but that it can be illuminated by the notion of real definition. x’s real definition specifies its essence, and consists of all the propositions that are true in virtue of the nature (or identity) of x, i.e. by all the propositions that result true when the operator ‘true in virtue of the nature of x’ is attached to them. These operators form a family: for any object or plurality of objects, there is an operator that specifies the propositions made true by their essences. Fine takes it that these operators should not be analyzed into a general notion of ‘true in virtue of’ on the one hand, and ‘the nature of’ on the other (Fine 1995b: 273).

Hale also endorses definitional essentialism (2013: 132), but his treatment of the locution ‘truth in virtue of the nature of X’ differs from Fine’s in that he takes it as expressing ‘a semantically complex relational property, composed of the relation expressed by “...is true in virtue of _” and the term “the nature of X”.’ (2013: 134) As Hale’s “...is true in virtue of _” denotes grounding, his view is that necessary truths are grounded in (but not reducible to: 2013: ch. 3) the essences.

According to Lowe, ‘the essence of something, X, is what X is, or what it is to be X’ (2008: 35). He would later identify this with ‘a key notion [...]’, pointed out and exploited by Fine himself’, which is real definition again (2013: 17), noting that ‘I could probably say all that I want to about my version of essentialism’ with the help of ‘locutions involving only sentential operators of the form “it is part of the essence of X that”—where “the essence of X” is not taken to make an independent contribution to the meaning of the operator’. So it seems that he would have identified his sentential operators with the Finean ones.

Then, we are dealing with a family of views endorsing the following theses:

1. Essences are given by real definitions.
2. Real definitions are expressed by ‘true in virtue of the nature of X’-operators (whether these are semantically complex or atomic).

---

4 See e.g. 2013: 116, 117, 131 n. 22, 155, 254.
Let the conjunction of (1) and (2) be Definitional Essentialism or DE. It is a view about how to understand essentialism. But it has been proposed that DE allows for an explanation of modality, to which we now turn.

III. EXPLAINING MODALITY

According to Fine (1994: 9),

[... F]ar from viewing essence as a special case of metaphysical necessity, we should view metaphysical necessity as a special case of essence. For each class of objects . . . will give rise to its own domain of necessary truths, the truths which flow from the nature of the objects in question. The metaphysically necessary truths can then be identified with the propositions which are true in virtue of the nature of all objects whatever.

The thought is that, in the explanatory order, first come the essences of the objects, and then the necessary truths emerge as those pertaining to the essence of ‘all objects whatever’. As we’ll see later, Fine proved a theorem in his logic of essence backing up this metaphysical claim.

Relatedly, Hale proposes a non-reductive, non-transmissive explanation of modality by the essences (2002; 2013, chapters 5 and 6):

The essentialist theory of modality claims that if it is metaphysically necessary that \( p \), there will be an explanation of the form: Necessarily \( p \), because, for some \( X_1, \ldots, X_n \), \( p \) is true in virtue of the natures of \( X_1, \ldots, X_n \) (i.e. \( \Box_{X_1,\ldots,X_n} p \)).

(2013: 155). The explanatory reading of Fine’s claim and Hale’s essentialist theory of modality share the general background idea that essences are to be put before modality in the order of metaphysical explanation.

Others have proposed that modality is grounded in the essences. Lowe defends that ‘essences are the ground of all metaphysical necessity and possibility’ (2008: 45; cf. 2013: 21), and Rosen suggests that ‘Fine might regard the irreducible facts of metaphysical modality as systematically grounded in the essences of things’ (2012: 121). If grounding is, or at least backs, an explanatory relation (in the sense that if \( x \) is grounded in \( y \), then \( x \) is explained by \( y \)), as is usually

5 Note that Lowe and others do not think of essences as separate objects—grasping the essence of \( x \) amounts to understanding what \( x \) is. In the rest of the paper, I will sometimes talk of ‘essences’ as if I was intending to refer to some objects; but this could always be rephrased in terms of essential truths, which is what I take as the core commitment of definitional essentialism.

6 Perhaps a stronger reading could be as a definition of metaphysical necessity by restriction on essentialist truth; or as a reduction of necessity to essence (cf. Cameron 2010: 333; Cameron re-considers the reductionist reading in 2012: 19). However, the weaker reading suffices for my purposes.

Kment (2014), ch. 6 voices his attraction to DE, although he aims to explain modality not only by essences but also by grounding connections.
assumed, then these suggestions can also be read as explanations of modality by essences.

The notion of explanation that I am using here is subject to these six constraints:

**Why** An explanation is an answer to a ‘why’ question—in our case, ‘Why are some propositions necessary rather than contingent, and some possible rather than impossible?’

**Acyclicity** That $x$ be explainable by $y$ requires that $y$ need not be in turn explained by $x$.

**Fundamentality** If $x$ is explanaible by $y$, then $y$ is more metaphysically fundamental than $x$.

**Ontic and factive** An explanation of $x$ by $y$ consists of truths about $y$—not about our concept or theory of $y$—that explain truths about $x$.9

**Structural** The explanation need not be of an object or property: necessity is not an object nor a property, but still, it is a legitimate philosophical task to try to explain it by something else.10

**Objective** An explanation of this kind does not constitutively depend on our epistemic powers: it might be that $x$ is explainable by $y$ even if no one knows about that; it might even be that we falsely though rationally believe that $x$ is not explainable by $y$. (That said, one should believe what the evidence indicates; for that is how we best avoid dogmatic conclusions.)

I think that this kind of explanation is usually assumed by metaphysicians, and that relations like (non-eliminative) reduction, truthmaking, ontological dependence, grounding, causation, supervenience, and constitution have many times been taken as metaphysically interesting because they allow us to phrase explanations of this kind.11 I won’t make further commitments as to whether this is captured by a specific notion of grounding, or of something else, or as to its exact logic; it suffices for my purposes that the explanation in question is (a) metaphysically important, and (b) subject to the constraints above. Stepping up from reduction, grounding, etc., to explanation, allows the arguments below be as general as to cover different specific proposals for the first-order explanatory relation between modality and essences.

---


9 I am aware that in the sciences it is legitimate to use idealizations—and thus, falsities—to give explanations. Perhaps idealization is also legitimate in metaphysical explanation; we will not delve into the issue here.

10 Sometimes, explanation is taken to be a relation among facts; here, however, I remain neutral on whether facts exist.

11 Cf. Schaffer: ‘there is a general kind of explanation which can be backed by causation of by grounding (inter alia), or in some cases by a mixture of the two’ (2012: 131). Fine thinks of ground as an operation rather than as a relation (2012: 43), but still accepts its connection to a relation of explanation (2012: Section 1.1).
I hope to have given enough by way of elucidation of explainability by. We can now frame the thesis that I want to argue against: that modality is explainable by essences, i.e.:

**MEE** Every modal truth is explainable by some essential truths.

The essentialist who wants to reduce, ground, etc., modality in the essences should believe in **MEE**, as (ontic) reduction and (full) grounding assume the kind of explanation I have just described. But given that a modal understanding of essence is hopeless as a starting point of the explanation, I will focus on **DE**, which, with **MEE**, gives us the thesis that modality is explained by definitional essence:

**MEDE** Every modal truth is explainable by some real definitions.

Of course, I assume that the real definitions **MEDE** is about should not include the real definition of metaphysical necessity itself, as this would clearly conflict with the **Fundamentality** constraint above.

IV. IS MODALITY EXPLAINABLE BY ESSENCE?

**MEDE** just tells us that there is an explanation of modality by the essences; it doesn’t tell us how is that explanation supposed to go. In this section, I examine and reject the currently available options.

**IV.1. Hale’s argument and non-transmissiveness**

Hale argued that ‘the supposition that a thing might have had a different nature immediately raises an obvious problem’ (2013: 133). If this were right, modality could be explained by essence, if the reason for the necessity of each necessary proposition were to be found in a thing having a given nature. Here’s Hale’s argument:

The supposition that α might have had a different nature is the supposition that it might not have been the case that Φα, and might have been that Φ′α instead. [...] This is equivalent to the supposition that for some β, it might have been the case that β = α∧¬Φβ∧Φ′β. But how could this be possibly true? Given that Φα tells us what it is for α to be the thing that it is, and that ¬Φβ, β lacks what it takes to be that thing, it must be that β ≠ α. In short, the supposition that a thing’s nature might have been different breaks down because it is indistinguishable from the supposition that something else lacks that nature.12

This argument fails.

Essences are what it is for something to be the what it is; it doesn’t follow that they are what it is for something to be what it must be. Differently put, I am claiming that, under DE, the essence of an object is what defines it, but that the metaphysical necessity of such a definition is a further posit. So, when Hale claims that ‘β lacks what it takes to be [α]’, he’s going a step too far. What is true is that β lacks (at the imagined situation) what it actually takes to be α; but a further premise is needed to infer from this that β lacks (at the imagined situation) what it takes to be α at the imagined situation. The premise that fills this gap is that essences are necessary, but it is a further premise—and of course, what Hale’s argument aims to establish.\textsuperscript{13}

Also problematic is Hale’s claim that ‘the supposition [1] that a thing’s nature might have been different breaks down because it is indistinguishable from the supposition [2] that something else lacks that nature’. Is it the claim that both suppositions are metaphysically indistinguishable? But I don’t think they are even semantically indistinguishable, if we let ‘α’ and ‘β’ be rigid designators. Then, supposition 1 is distinguishable from supposition 2 by noting that 1 embodies the stipulation that α = β; while 2 embodies the stipulation that α \neq β. Supposition 1 is absurd only if it is assumed that essences are necessary; but again, this is what Hale’s argument purports to show.

Hale aimed to establish the necessity of essential truth so as to be able to give a non-transmissive explanation of necessity in terms of essential truth. The necessity of a proposition is non-transmissively explained by another proposition, q, iff q is necessary, but its necessity is not appealed to in the explanation (2013:131).

Consider the following example. What, according to Hale, explains that \((A \wedge B) \Rightarrow A\) is the plain fact that by its very nature, conjunction is such-and-such a function, not its necessitation (2013:132). The fact is indeed necessary; but its necessity does not enter the explanation. But how can that be? I fear that the necessity of logical facts may corrupt our intuitions here, so let’s use another example.

Suppose that it turns out that a certain waste bin is essentially made out of iron. I claim that this fact stops short of explaining why it is that the waste bin is necessarily made out of iron. If the metaphysical necessity of real definitions is a modal posit, then the necessity of the waste bin’s constitution is not explained just by the fact that the waste bin’s definition includes its constitution: the modal posit is needed. Of course, the explanation takes off if appeal is made to the necessity of constitution; but that makes the explanation transmissive.

\textsuperscript{13} A referee pointed out that even someone who held essences to be contingent, could grant that x being essentially \(\Phi\) implies x being necessarily \(\Phi\). Perhaps the conjunction of contingent essentialism with the entailment of necessary truth from essential truth is indeed coherent. But I am not arguing against its coherency; I am asking what reasons are there to believe that essence does explain necessity. Hale offered an argument to that effect, and I have argued that his argument is fallacious.
That is, once essence is understood as real definition, a conceptual gap opens: a gap between a being \(F\) by its very nature, and a being \(F\) necessarily. The existence of this conceptual gap entails that a being essentially \(F\) does not by itself explain why it is that \(a\) is necessarily \(F\). And my worry is that essentialist explanations of modality seem to be non-transmissive only because the modal principle of the necessity of real definitions is implicitly assumed. As will be argued next, this principle may seem obvious, but nonetheless it is a modal postulate.

IV.2. *Ab initio necessity vs the gap*

Some philosophers may want to claim that ‘all essential truths are necessary’ is analytically true. But I suspect that much of the motivation for the analyticity claim is to be found in the supposed *obviousness* of the necessity of essential truth. If so, analyticity is a property stronger than what I call the ‘*ab initio necessity*’ of essential truth: the obviousness of the necessity of essential truth that whoever has grasped the notion of *essence* will allegedly feel. This obviousness, as I am thinking of it, would need no further connecting principle or theory—it would be obvious that essential truth is ‘directly’ or ‘from the very beginning’ necessary.

If essential truths were *ab initio* necessary, Hale’s argument above would be vindicated, and perhaps the explanation of modality by essence would mostly rest on that. But I don’t think that essential truths are *ab initio* necessary; I think that we need a theory of essence in order to give a reasoned verdict. (As Lowe wrote, a ‘semantic analysis’ of essentialist expressions ‘*is no doubt an exercise that should be undertaken at some stage in a full account of what I am calling serious essentialism*’; 2008: 35, my emphasis).

The essentialist might claim that the introductory examples of real definitions ‘wear their necessity on their sleeve’: we see that they are necessary just by understanding them. But I don’t think that the usual examples—like ‘Socrates is a man’ or ‘a natural number is a finite ordinal’—are like that. I think that they are aptly understood as saying what the object is—and this is, *furthermore*, necessary. Under this second view, the necessity of real definitions does not follow merely from their being real definitions: it is a *further* posit, a separate modal fact: a fact not just about essences, but also about modality. And this second view is compatible with whatever intuitions are pumped by the bare examples. When I am told that Socrates is essentially human, I understand that humanity is part of the nature of Socrates. If am questioned whether, given that Socrates is essentially human, he could have been not human, I feel the urge to answer in the negative—but this time, the intuition flows not only from my grasping what it means that he be essentially such and such, but also from my grasping what it means for something to be necessary. That is, my intuition is a belief in a *modal hypothesis*: one about the extent of necessity.
So, the choice between \textit{ab initio} necessity and the second view is underdetermined. Equivalently: one can understand the notion of real definition and still see a gap between saying what an object \textit{is}, and saying what the object is \textit{necessarily}. And if essential truth is not \textit{ab initio} necessary, then, it’s not a conceptual mistake to question the necessity of essential truth: its necessity is not seen just by aptly grasping the relevant concepts. One needs a principle or a theory that entails so.

\textbf{IV.3. Logic?}

Perhaps essential truth is not \textit{ab initio} necessary, but something very close to it: it is necessary \textit{by logic}. If so, then perhaps that \( p \) is necessary is explainable by some real definitions: because some object is by nature such that \( p \), and then logic entails that \( p \) is necessary. Consider this argument:

\begin{enumerate}
  \item Nominal definitions have the status of axioms.
  \item So, every logical system (with a definable notion of theoremhood) proves the definitions made in its language.
  \item The rule of Necessitation (if \( \vdash A \) then \( \vdash \Box A \)) makes those definitions metaphysically necessary.
  \item But real definitions are like nominal definitions.
  \item Therefore, real definitions are metaphysically necessary.\footnote{I am deeply thankful to Alessandro Torza for suggesting that I should consider this potential line of response.}
\end{enumerate}

I concede the analogical claim (4) as following from \textbf{DE}, and it might be argued that (1) and (2) are true by a general conception of logic. But what justifies (3)? Necessitation fails in some modal logics: the \textit{non-normal} ones. Why are these not considered in the argument? It must be because it is assumed that the logic of metaphysical necessity has to be normal (say that ‘metaphysical modality is normal’).\footnote{Strictly speaking, this needn’t be \textit{assumed}; it might be held on \textit{(for example) abductive grounds. No matter—What is problematic for the essentialist here is the thesis itself.} So, even if some object is by nature such that \( p \), pure logic is not enough to infer that \( p \) is necessary: one has to assume that metaphysical modality is normal. This assumption would make the purported explanation clash with the \textbf{Acyclicity} constraint.

It will not do to assume that the logic of metaphysical necessity is non-normal; it will not do to assume that the logic in (3) systematizes any other modality. Or will it? Perhaps the logic of \textit{logical} modality is normal, and perhaps logical necessity entails metaphysical necessity? So every essential truth is logically necessary [suitably modifying (3)], and this entails that essential truth is metaphysically necessary.

But the assumption that logical necessity entails metaphysical necessity is, if not Necessitation itself, very close to it. Why is it sound to infer, from \( p \)’s logical
necessity—in the formal mode: from \( p \) being a theorem in a sound system of logic—that \( p \) is metaphysically necessary? This is not merely a principle about logic. It is a principle about logical and metaphysical necessity. Once again it is clear that (3) embodies an assumption about metaphysical necessity, one left unexplained by the essences.\(^{16}\)

I conclude that, if real definitions are necessary, this has to follow from a metaphysical theory, as it does not from either logic or by mere grasp of the notion of essence.

\**IV.4. Finean semantics for essence**

Perhaps, the explanation flows from the Finean framework.

Remember that Fine identifies the necessary truths with ‘the propositions which are true in virtue of the nature of all objects’. But what does it mean to say that a proposition is \textit{true in virtue of the nature of something}? And how can we be sure that \textit{all and only} the necessary propositions are the ones true in virtue of the nature of all objects? And what does ‘\textit{all objects}’ mean in this context?

Let us start with the last question. ‘\textit{All objects}’ should mean something more encompassing than ‘\textit{all objects that actually exist}’. As we will see, Fine’s semantics entails that: possibly, \( p \) iff it is not true in virtue of the nature of all objects that not-\( p \)—iff there is an object among \textit{all the objects} such that it is not true in virtue of its nature that not-\( p \). So, we just have to find a proposition \( p \) such that (1) \( p \) is possible, but (2) it is of the essence of all actual objects that \( p \) is false. Some metaphysical theories entail (1) and (2). It might turn out that both nominalism and a ‘laws-first’ theory of material objects are true: there are only material objects and these are essentially structured by the laws of nature. If so, although (1) there could have been objects evolving under different laws, (2) that they don’t exist is part of the essence of all actually existing objects—as it is part of \textit{their} essence that the actual laws hold.

Then, ‘\textit{all objects}’ in Fine’s claim has to be a quantifier not restricted to the actually existing objects—and indeed it is, in Fine’s logic and semantics, a \textit{possibilist} quantifier (1995a; 2000). It is hard to escape the conclusion that Fine’s framework introduces primitive modal ideology from this point.\(^{17}\)

The objection-recipe given by (1)-(2) above assumes that there might have been something different to everything that actually exists; but some reject this assumption.\(^{18}\) But more problems are forthcoming, turning now to the second

\(^{16}\) Cameron (2012): 15-7 has objected to this argumentative strategy—showing that unexplained modal truths or ideology are presupposed—against reductionist proposals about modality. I have strong doubts about his objections, but responding to them would take us far afield. Here, I cannot do much more than rely on the constraints on metaphysical explanation laid out in Section III.

\(^{17}\) For related worries, see Teitel (2018: 20–23).

\(^{18}\) Williamson 2013.
question: How can we be sure that all and only the necessary propositions are the ones true in virtue of the nature of all objects?

Here’s a possible reply: if essential truths are ab initio necessary, ‘any essentialist attribution will give rise to a necessary truth’ (Fine 1994: 8). I have already rejected that by merely understanding the usual examples of essentialist claims we see their necessity. We want a theory. Fine has given one, in the form of a logic and model theory of essence (1995a, 2000). Which leads us to our first question: what does it mean to say that a proposition is true in virtue of the nature of something?

We must tread carefully here. Fine’s model theory is the best we have for a semantic analysis of essentialist sentences.19 But we should not assume that it does, or should, give a reductive analysis of the locution ‘true in virtue of the nature of’. The assumption, rather, is that the semantics should constrain the theoretical architecture of the essentialist framework, allowing us to understand its fundamental concepts when not by analysis, then by synthesis: by the way they relate to other, well-understood, concepts.

In (2000: 548), Fine defines the truth at a world of a proposition of the form $\Box_F(A)$, which intuitively means that $A$ is true in virtue of the nature of some (or all) $F$s as:

**Essential Truth** $w \models \Box_F(A)$ iff:

- (a) $[A]^M \subseteq c(w_F)$, and
- (b) for all $v \in W$: if $F_w \subseteq I_v$, then $v \models A$.

Unpacking this definition, and saying that $y$ supports $x$ whenever $x$ depends on $y (x \succeq y)$, **Essential Truth** amounts to this: $\Box_F(A)$ is true at $w$ (in model $M$) iff each object mentioned in $A$ (according to $M$) is either an $F$-at-$w$ or supports some $F$-at-$w$, and $A$ is true at any world that contains the $F$s-at-$w$.20 (Later, we’ll take issue with this ‘containing’.)

It now seems that the only concept involved in **Essential Truth** that remains to be explained is that of dependence. Fine defines dependence in terms of essentiality (1995a: 275); but it would be unfair to press a circularity charge;

---

19 Other logics of essence are Marcos (2005) and Giordani (2014), but neither is an acceptable framework for an explanation of modality: we learn from Section 1 of Marcos’s paper that his essentialist operator $\circ$ is to be interpreted as necessity conditional on truth; and we learn from Section 2 of Giordani’s paper that his system includes a primitive predicate, $\mathbf{Ex}$, the intended interpretation of which is possible existence. Therefore, the intended interpretation of both alternative approaches is unexplainedly modal from the start, and no other interpretation of them that could be relevant to our topic is immediate. Correia (2000) has developed a propositional version of Fine’s logic of essence, but its basic model theory is also subject to the difficulties I develop below.

20 As an exercise for the reader, it might be instructive to see how does the famed asymmetry (Fine 1994) between Socrates and his singleton $\{S\}$—that it is essential to $\{S\}$, but not to Socrates, that Socrates $\in \{S\}$—flows from **Essential Truth**.
we were never promised a reductive definition of *essence*—it is the starting, primitive point.

It remains to show that all and only the metaphysically necessary truths are essential truths. To do so, Fine assumes that the correct logic of metaphysical necessity is the quantified version of S5 with constant domain, S5π. (Perhaps analogous theorems could be proved with respect to different quantified modal logics—perhaps.) Then, Fine proves that S5π-theoremhood is a special case of E5+-theoremhood, where E5+ includes the axioms of first-order logic, the axioms for $x \geq y$, the $\Box_F$-operators, and the predicate abstractor $\lambda$, and the Domain axioms. The first one of these says that the existence of every possible object is compatible with the nature of every other possible object; and the second says that it is true in virtue of the nature of all the possible objects that they are all the possible objects (1995a: 250).

Let’s be a little more specific. Fine defines ‘$\forall$’ as $\lambda x[\Pi x(x = x)]$, where ‘$\Pi$’ is the possibilist universal quantifier. So ‘$\forall$’ is the property had by every possible object, and ‘$\Box \forall$’ is the ‘maximal necessity’: truth in virtue of the nature of all possible objects (1995a: 246). Fine proves that E5+ ‘yields’ S5π in the sense that, first, ‘Any formula $A$ of S5π may be translated into a formula $A'$ of [E5+] by replacing each occurrence of $\Box$ with $\Box \forall$', and second, ‘For any theorem $A$ of S5π, $A'$ is a theorem of E5+’ (p. 267).

This indeed shows that S5π-necessity is a special case of E5+-essentialist truth. But it does not explain necessity by essentialist truth, as the ‘reducing’ logic is replete with unexplained modal notions. If they cannot be explained, MEDE is false under this option, because the semantics makes essence require explanatorily previous modal truths.

First, as already underlined, E5+ includes possibilist quantifiers as primitives, which seems to take possible objects for granted. A reply is that, given that the Domain axioms are assumed, the resulting logic E5+ corresponds to models with constant domains, i.e. the actual objects are all the possible objects. The falsity of ‘there could be something different to all that actually exists’ follows immediately. However, (1) this assumes controversial theses, which are (2) essentialist and modal theses put ‘by hand’, not explained.

Secondly, let’s get back to Essential Truth. It is defined partly in terms of the schematic condition ‘$x \in I_w$’ (*The Domain Condition*), where $I_w$ is the domain of the world $w$. Fine stipulated that ‘The presence of an object in a world is not taken to guarantee its existence but merely its possibility’ (2000: 543; my emphasis). So at least the informal gloss of a crucial notion in Fine’s semantics also assumes modal ideology: the notion of an object being possible at a world.

The source of the problem is not the notion of world; as it is just the notion of a complete, maximal, reality—or a representation thereof. Even the notion of a possible world need not be taken to be a fundamentally modal one, if Lewis is right that his worlds reduce modality (Lewis 1986). But if the essentialist
were to use Lewis’ cosmology, she would ultimately explain modality not by essences, but by Lewis’ cosmology.

While I don’t think that we can completely sidestep issues in this vicinity, we can concede the essentialist her worlds, for, as we’ll see below, she is able to logically construct them—if she assumes modal facts. That’s where the trouble really lies.

Fine’s claim that ‘the presence of an object in a world is not taken to guarantee its existence but merely its possibility’, taken at face value, entails that The Domain Condition is modal: it is satisfied by a only if a is possible at w. One natural way to try to do without modality is to identify \( I_w \) with the set of objects possible at w, but with the set of objects existing at w. And indeed Fine does so (2000: 550). However, this is done for purely technical reasons: to simplify the proofs. It would be a mistake to read off a metaphysical theory of essence from that simplification, as Fine himself noted (2000: 550).

It is against the modal-existential view of essentialist claims that Fine forcefully argued against in 1994, because (i) it presupposes that essence is a special case of necessity, but (ii) essence is much more fine-grained than necessity.

Thus, the modal-existential interpretation of the formalism should be taken as an artefact of the formal model, not as a feature of the metaphysical reality modelled by the essentialist logic. But then we are back to the drawing board: How should we understand The Domain Condition?

Fine writes (2000: 543; my emphasis):

The basic idea behind the semantics is that a statement should be taken to be true in virtue of the nature of certain objects if it is true in any world compatible with the nature of those objects. We shall make the simplifying assumption that each world is compatible with the nature of all and only those objects that it contains.

But what does it mean to say that a world w is compatible with the nature of the objects in \( I_w \)? The meaning of ‘compatibility’ that immediately suggests itself is a modal one: two propositions are compatible if they can be jointly true; two objects are compatible if they can coexist.

Perhaps, the essentialist can define her own notion of compatibility. Say that x is essentially compatible with a world w if it is not of the essence of the objects in \( I_w \) that x doesn’t exist (this way, the interpretation of Fine’s logic would be given in that very system; but we’re used to assuming theories at their own meta-level.) A familiar recursive construction defines each \( I_w \) as closed under essential compatibility.

However, this will be useless for the essentialist unless she makes two assumptions: (1) that all the objects in the domains are metaphysically possible, and (2) that each metaphysically possible object is in one domain or other. Without

---

21 I thank Prof. Fine and an anonymous referee for suggesting the following line of response.
(1) in place, nothing prevents some impossible object ending up in a domain.\textsuperscript{22} Without (2) in place, nothing prevents some possible object not being an element of any domain. So, without (1) and (2) and even assuming the essentialist compatibility criterion, nothing prevents Fine’s models to either contain impossible objects or not contain some possible objects. But these models cannot be used to explain modality: for the first case, truth in virtue of the nature of an impossible object would explain necessary truth, and for the second, some possible truth would remain unexplained.

The Finean might object that the assumptions are not, in the end, required. ‘Even if some Finean domains contain impossible objects, and some lack some possible objects, set theory entails that there are domains containing all and only the possible objects. Pick out the latter’, says the Finean, ‘and with them, give the intended model of the logic; explain modality with it’.

But set theory only entails that, for every definable formula of our language, there is a set of the objects satisfying the formula.\textsuperscript{23} The Finean needs it to be the case that there is a formula satisfied by all and only the metaphysically possible objects. This is not entailed by set theory alone—and no non-modal principle that will do the work comes to mind. Equivalently put, the semantics of the logic of essence can give the explanation only given the modal assumption that there is a set containing all and only the possible objects. This does not follow from set theory alone and constitutes an assumption not explained by the essences.

In short: there are various ways to construct models for Fine’s essentialist logic; some contain domains containing impossible objects, some leave some possible objects out of every domain. But for the intended model to exist, the assumption that certain domains contain all and only the possible objects has to be in place. This modal assumption is not explained within Fine’s semantics.

Therefore, neither the existential interpretation of the formal semantics nor the appeal to either a modal or an essentialist notion of compatibility, allows us to use Fine’s semantics to explain necessity by the essences.

One option is to go anti-realist, and claim that conceivability, imaginability, or some other mind-confferred feature is what it takes for an object to be a member of a domain [in the lines of Sidelle (1989) or Sveinsdóttir (2008).] For

\textsuperscript{22} Objection: impossible objects do not exist; so they can’t ‘end up’ in any set. Rejoinder: Neither do merely possible objects, but they are needed to give the intended model. Objection: But in Fine’s logic there are no merely possible objects: all the possible objects are the objects that exist. This follows from the Domain axioms. Rejoinder: But these axioms are about the possible objects. Take this unexplained modality out and the second axiom reads: ‘it is true in virtue of the nature of all the objects in the domains that they are all the objects in the domains’. This does not entail that all the possible objects are the objects that exist. So, the Finean manages to define her models as containing just possible objects only by relying on an unexplained modality.

\textsuperscript{23} This glosses over the specific constraints that each theory will enforce to avoid contradiction. In ZFC, this is done by the axiom schema of comprehension.
example, one could claim that The Domain Condition is satisfied by any object only if its existence at \( w \) is conceivable.

But this would spread mind-dependency all the way up to the notion of essentiality—and, given the essentialist project here considered, all the way up to metaphysical necessity. But surely one would prefer to be argued to anti-realism about necessity, rather than being thrown into it by problems intrinsic to essentialist reductivism. Be that as it may, there’s also the worry that it better be that being conceivable need not be understood as being metaphorically possibly conceived (Bealer 2002: 75–6), for then we have, again, smuggled a metaphysico-modal notion.

We might try to resort to logical modality, for in a sense it seems weaker than both actual existence and metaphysical possibility: the logically possible, it is widely held, need not be metaphysically possible—and so, logical modality is not plausibly defined in terms of metaphysical modality. Further, on most accounts, logical modality is mind-independent, and thus presents no risk of spreading anti-realism.

Say that \( x, y \) are objectually-logically compatible (\( OLC_{xy} \)) if the conjunction of their real definitions does not entail a logical contradiction. Then, \( OLC \) is symmetric, but not transitive; it is reflexive when restricted to the logically possible objects: those complementary to the logically impossible objects: those logically incompatible with themselves (i.e., their real definition conjoined with itself entails a logical contradiction), and therefore with any other object.

If the logically possible objects are countable, they can be enumerated \((x_1, x_2, \ldots)\) and then a Lindenbaum-style construction can define domains as maximal, \( OLC \)-consistent sets:

1. Take an \( I_w \) such that \( \forall x, y \in I_w: OLC_{xy} \). Let this be \( I^0_w \).
2. Let \( I^i_w = I^{i-1}_w \cup \{x_{i+1}\} \) if \( \forall y \in I^{i-1}_w: OLC_{x_{i+1}y} \), and \( I^{i-1}_w = I^{i-1}_w \) if not.
3. Let \( I^*_w = \bigcup_{i=0}^{\infty}(I^i_w) \).

It is easily seen that each \( I^*_w \): (i) contains only objects that are \( OLC \) with each other, and (ii) is such that no further object can be added to it while preserving \( OLC \). 24 Given all this, \( a \) satisfies The Domain Condition if\( \text{iff} \) either there is some \( x \in I^*_w \) such that \( x \geq a \), or if \( OLC_{ax} \) for every \( x \in I^*_w \).

Unfortunately, appealing to logical necessity won’t help the Finean. Not everything that is logically possible is thereby metaphysically possible [Pruss (2015) proves this from Gödel’s second incompleteness theorem.] So even if two objects are objectually logically compatible, it does not follow that they are metaphysically compatible: it does not follow that their essences do not entail

24 For the domains \( I^*_w \) to be closed—as Fine required—under dependence (\( \geq \)), the essentialist also needs the non-trivial assumption that \( x \geq y \) entails that \( OLC_{yz} \) for all \( z: OLC_{xz} \). Otherwise, there will be \( x, y \) and \( w \) such that: \( x \in I^*_w, x \geq y, \) but \( y \not\in I^*_w \) because \( \neg OLC_{yz} \) for some \( z \in I^*_w \).
a contradiction given the metaphysically necessary truths. How could that happen?

Perhaps there are metaphysically-impossible-but-logically-consistent essences. Or even if there are no metaphysically impossible essences, perhaps logical compatibility does not preserve metaphysical compatibility. Let’s see.

First, as we saw above, the only way to prevent assigning objects with logically possible but metaphysically impossible essences to the domains is to place an unexplained modal criterion. With no such criterion in place, let the TimeMaster be someone who can make time at the actual world be either a line-segment or a loop at will; or take Leibnizian God, who can will something to be metaphysically possible or impossible. They have clear, logically consistent definitions, but they are metaphysically impossible. Therefore, even if some \(x\) is OLC with TimeMaster, that does not entail that \(x\) is metaphysically compatible with TimeMaster—because nothing could coexist with it, given that it could not exist.

Secondly, it follows independently from a variety of metaphysical hypotheses that some objects are OLC, but not metaphysically compatible—even though each is metaphysically possible. For example: causal overdetermination is not contradictory, but if you believe it to be metaphysically impossible, consider the causal process of \(c\) producing \(e\), and the process of \(c^*\) producing \(e\) (where \(c\) and \(c^*\) are different possible causes of \(e\)). They are logically compatible—but not metaphysically so.

---

25 Kment assumes this principle: ‘it’s essential to \(e\) that \(P\) logically entails the existence of \(e\) and of every entity mentioned in \(P\); commenting that he ‘can’t do much better to justify [it] than to say that it seems to me to be obviously true’ (2014: 157). If this principle were true, it would favour the essentialist here, as no impossible object would have an essence. But I think that it is highly non-obvious, for two reasons. First: different conceptions of essential properties are more or less friendly to it. If essences are necessarily existent properties, even the essences of non-existent objects will exist (cf. Plantinga 1976, section II.5, who thinks that even the essences of contingent objects are necessarily existing.) If (more on the lines of DE) essences are sets of propositions specifying real definitions, and if propositions do not depend on the objects they are about, even the essences of contingent objects will be necessarily existing. These views are neither obviously correct nor obviously absurd. Or perhaps essences are not further objects: there are only true or false sentences attributing real definitions to objects. But then, the essentialist theory needs much more detail in order to entail that sentences attributing real definitions to non-existing objects are all false. Second, the conception of essentiality as existentially neutral goes back at least to Aquinas (Mondin 1975; ch. 4), and was endorsed by metaphysical giants like Leibniz or Descartes (McDaniel 2017: ch. 9). Indeed, more radically, people like Spinoza (and myself) think that essentiality is also possibility-neutral: there are true ‘it’s essential to \(e\) that \(P\)’ claims, where \(e\)’s existence is metaphysically impossible. This explains the literal truth of claims that something by its very nature cannot exist, as: ‘Russell’s paradoxical set, by its very nature, cannot exist’, or ‘The TimeMaster’s essence makes its existence impossible’.  

26 For related objections, see Wildman (2018: Section 2). These objections do not work if the first Domain axiom is presupposed and compatibility is read essentially or modally; but we are interpreting compatibility as OLC here.
Thirdly, mathematical theories are metaphysically necessary (or so, at least, it is widely believed). Thus, ZFCU (standard set theory with *urelemente*) must be so. Thus, as a matter of metaphysical necessity, Socrates’ existence entails {Socrates}’s existence. But not as a matter of logical necessity—if an appropriate kind of logicism is false, ZFCU is not logically necessary, as it contains the non-logical primitives ‘∈’ and ‘Urelement’. Therefore, Socrates’ existence is logically, but not metaphysically, compatible with {Socrates}’s non-existence.

**Moral:** *objectual-logical compatibility does not entail metaphysical compatibility.*

But then the essentialist can’t use logical compatibility to secure material adequacy, as there can be a and b such that (i) both satisfy The Domain Condition for *w* and, therefore, both are in *I*; but (ii) they can’t metaphysically possibly coexist. So *w* would be a metaphysically impossible world. And these worlds cannot be used to explain possible truth. So, logical modality is of no use in defining The Domain Condition.

We want to use Essential Truth for a semantics of essentialist discourse; doing so requires us to steer a way between the modal-existential account of essence, and abandoning MEDE. Neither existence nor metaphysical, epistemic or logical modality were useful for this. There might be more options; but don’t hold your breath: if there’s any ontological status different from these, metaphysicians have been overlooking it for quite a long time.27

**IV.5. Real definitions and grounding**

Let’s get back to DE. It understands essence through real definition. Perhaps a theory of real definitions uncommitted to the Finean operators—and thus to their logic and semantics—will help to explain modality by the essences.

Rosen (2015); Fine (2015); and Correia (2017) have offered approaches to real definition that do not turn on Fine’s logic of essence. As Correia (Section 7) has argued that his accounts supersedes Rosen’s and Fine’s, let us focus on it.

Correia takes real definitions to be of the form ‘To be *F* is to be *G*’. It is important to differentiate the ‘is’ of real definition from the ‘is’ of what he and others have called ‘generalised identities’ (as in: ‘For Carlos to instantiate the property of writing is for Carlos to write’). Let us use ‘≡’ for the former and ‘i≡’ for the latter. It’s important to make such a distinction as real definitions are asymmetric—if to be *F* =D to be *G*, then the converse does not hold; while generalised identities are symmetric: if to be *F* ≡i to be *G*, then the converse does hold.

---

27 Alessandro Torza suggested to me that it may be worthwhile exploring whether *Meinongian subsistence* has what it takes for this role. That obviously goes beyond what I can offer in the present discussion, but let me state upfront that I doubt whether an explanation of modality by essences and a non-existential kind of being (subsistence) gives all, or perhaps even many, of the benefits (like a reduction in ideological commitments) that were sought in the first place.
However, Correia believes that generalised identities and real definitions are importantly connected:

**Identity Principle** If \( F \) to be \( F \) to be \( G \), then to be \( F \) to be \( G \).

In order to block a paradox that could be derived from **IP**, Leibniz’s Law for \( \equiv \), and the irreflexivity of \( \equiv \), Correia is pressured to hold that the ‘is’ of real definition induces opaque contexts. Which leads to this further principle:

**Priority Principle** If to be \( F \) to be \( G \), then being \( G \) is metaphysically prior to being \( F \).

Either of two notions of priority can be plugged into **PP** to make it plausible: representational generic grounding and being more joint carving than.

Generic grounding is a relation between properties or concepts; while representational grounding is a relation between propositions or concepts. So, representational generic grounding is a relation between concepts.\(^{28}\) In this sense, \( p \) grounds \( \neg \neg p \), even though they may be said to be the same facts and grounding is irreflexive.

**Being more joint carving than** is related to Lewis (1983) naturalness and to Sider’s (2012) structure. Roughly, a concept \( F \) is more joint carving than a concept \( G \) whenever the first stands for a more fundamental aspect of reality than the second.

Reading **PP** with either representational generic grounding or relative joint-carpvingness, it secures both the irreflexive and the opacity-inducing nature of \( \equiv \): under either reading, it is clear that if \( F \) is metaphysically prior to \( G \), then the converse is false; and both readings appeal to concepts, but clearly, co-referring concepts can fail to obey Leibniz’s Law (viz. water and \( H_2O \)).

With all this in place, Correia’s account of real definition is this (p. 60):

**RD** To be \( F \) to be \( G \) iff (i) to be \( F \) to be \( G \), and (ii) being \( G \) is metaphysically prior to being \( F \).

Real definitions of things—‘To be Socrates is to be the son of Sophroniscus and Phaenarete’, say—are accounted for, when phrased in the manner of ‘To be identical to Socrates \( \equiv \) to be the son of Sophroniscus and Phaenarete.

Having our best-developed theory of real definitions at hand, can we use it to explain modality? Suppose that it is necessary that Socrates is the son of Sophroniscus and Phaenarete. Why is this so? Perhaps because (i) to be identical to Socrates \( \equiv \) to be the son of Sophroniscus and Phaenarete, and (ii) being the son of Sophroniscus and Phaenarete is metaphysically prior to being identical to Socrates.

But this won’t do.

---

\(^{28}\) This assumes understanding grounding as a relation; while it might be understood as a sentential operator. My arguments below can be cast on either understanding.
Consider RD’s first conjunct \((F_i = G)\). How could we explain the necessity of essential truth from this? Well, perhaps, if \(F_i = G\), then that is necessary. And so, as every real definition entails a claim like ‘\(F_i = G\)’, then, as these claims are necessary, real definitions are also necessary.

But why is \(F_i = G\) necessary? Well, it is supposed to be an identity of sorts; and the classical identity relation is necessary. So, if there is a certain analogy between these relations—or if \(\equiv_i\) is indeed a generalization of = that preserves the modal profile of =—then \(\equiv_i\) is necessary.

But why is = necessary? I can think of only two ways to respond. First, because of modal logic: it is a theorem that if \(a = b\), then \(\Box a = b\) (for any rigid designators \(a, b\)). But there are legions of modal-logical systems; that something is a theorem of one of them—or even of all—says nothing about metaphysical necessity. What is needed is the further claim that the theorem belongs to the correct system of modal logic: the one that indeed captures the logical properties of metaphysical necessity. But then, this way of justifying the necessity of identity assumes that metaphysical necessity has certain logical properties. And this assumption is left unexplained by the essences.

The other way is to cut the middle man out and claim that identity itself is necessary. But how is this not a modal presupposition? To assume that a certain relation is necessary is to make a modal assumption. But perhaps = is necessary because of its essence; so, ultimately, the necessity of real definitions indeed is explained in essentialist terms: it is explained by the essence of =, which, in turn, makes \(\equiv_i\) necessary, which in turn makes real definitions necessary.

However, if it is of the essence that = is necessary, then, according to RD:

\[ = \equiv_i R, \]

where \(R\) is a more fundamental relation, and \(R\) is necessary. But what relation is that? And why is it necessary?

Perhaps the necessity of \(\equiv_i\) is not to be explained by the necessity of =. Perhaps \(\equiv_i\) itself is necessary. Not because of its essence, of course: again from RD, we would get that

\[ \text{to be } \equiv_i \text{ = to be } R, \]

with an \(R\) for which again we have no idea what it is nor why is it necessary.

But perhaps \(\equiv_i\) is ab initio necessary. Indeed, theorists of the notion have suggested that it is closely connected to metaphysical modality.29

---

29 Dorr (2016) says that \(\equiv_i\) is necessary, and even that the fact that \(F_i = G\) explains the fact that necessarily all \(F_i\) are \(G_i\). Rayo (2013: 50) comments that the connections between \(\equiv_i\) and metaphysical necessity shed light on both notions, and shows (ch. 5) how far the connection can be pushed. For critical discussion, see Linnebo (2014).
Something analogous could be said about RD’s second conjunct in its grounding version: perhaps grounding is *ab initio* necessary. Indeed, many grounding theorists have suggested that what is grounded is necessarily entailed by what grounds it,\(^{30}\) or that the principles governing grounding are ‘metaphysical laws’ which define what is metaphysically possible.\(^{31}\)

I have put forward my doubts about *ab initio* necessity in Section IV.2. I’d like to press the point with respect to generalized identities and grounding. Why are these relations necessary? It isn’t a part of their real definitions, as RD then entails that they are ‘generalizedly identical’ to other, more fundamental relations, which are necessary—but we don’t know which ones are those nor why are they necessary.

So the *ab initio* necessity of these relations does not follow from their essences. What explains it, then? If it is posited as a primitive fact, it’s hard to convince oneself that this is somehow different from positing primitive modality.\(^{32}\) To be clear: I am not saying that grounding is not necessary—I am saying that, if it is, *that* is a *posit*, not something that is *explained* by grounding, nor by essence.

Lastly, turn to RD’s second conjunct in its relative joint-carvingness version. Here, the problem is that relative joint-carvingness ($<_{JC}$) does not, by itself, entail anything about other possibilities. In particular, $F <_{JC} G$ does not entail that $\Box (F <_{JC} G)$. Sider (2012) put no modal constraints on joint-carvingness—and indeed defended that modality is non-fundamental (ch. 12). Be that as it may, it just isn’t obvious that joint-carvingness by itself has modal implications.

Now, if a property is more natural in Lewis’ sense then this is necessary. However, the necessity of naturalness is too bound up with the specifics of Lewis’ system.\(^{33}\) Apart from that, if according to RD (i) $F \equiv_i G$, it’s quite difficult to see how could $F$ be more (or less) natural than $G$, as the naturalness scale is an order on the properties themselves, not on our representations of them.

V. CONCLUSION

I have examined five attempts to explain modality by essence, and have found that they either assume unexplained modal principles or ideology, or fail to

---


\(^{31}\) On the notion of metaphysical laws, see e.g. Glazier (2016); Kment (2014); Rosen (2012): 133; Sider 2012: Section 12.4.

\(^{32}\) Trogdon (2013) argues that grounding is necessary, noting that arguments have been lacking with respect to this issue. Unfortunately, his argument outright assumes that essential truths are necessary (p. 467), which is what we wanted to *show* with the help of grounding. Zylstra (2018) also argues for an analogous of the necessity of grounding in his framework. However, he reduces grounding to essentialist notions, while we’re trying to use grounding to illuminate real definitions. Others have also tried to understand, even if not define, grounding in terms of essence: Audi (2012): 108; Fine (2012): 76; Rosen (2012): 131.

\(^{33}\) Dorr & Hawthorne 2013: 31-33.
comply with the constraints on metaphysical explanation. Clearly, this is not a proof; these are just the most natural ways to develop the explanation. But I conclude that our present evidence favours rejecting the hypothesis that modality is explainable by essence.\footnote{The research for this paper was supported by PAPIIT IN401016 (Individuación en semántica, metafísica y epistemología) and by a CONACyT graduate grant. I have received feedback from many colleagues; specifically, I want to thank Víctor Cantero, Luis Estrada González, Sergio Gallegos, Kit Fine, Martin Glazier, Melahuaç Hernández, Moisés Maclas Bustos, Erick Llamas, Alessandro Torza, Lourdes Valdivia, and Jennifer Wang; participants of the Sabatine discussion group at the Philosophy and Literature Department (FFyL) at UNAM, México; participants of the Metaphysics Seminar at the Philosophical Research Institute (IIF) at UNAM; participants at the 2016 Graduate Conference at IIIF; and anonymous referees for invaluable comments.}

REFERENCES


34 The research for this paper was supported by PAPIIT IN401016 (Individuación en semántica, metafísica y epistemología) and by a CONACyT graduate grant. I have received feedback from many colleagues; specifically, I want to thank Víctor Cantero, Luis Estrada González, Sergio Gallegos, Kit Fine, Martin Glazier, Melahuaç Hernández, Moisés Maclas Bustos, Erick Llamas, Alessandro Torza, Lourdes Valdivia, and Jennifer Wang; participants of the Sabatine discussion group at the Philosophy and Literature Department (FFyL) at UNAM, México; participants of the Metaphysics Seminar at the Philosophical Research Institute (IIF) at UNAM; participants at the 2016 Graduate Conference at IIIF; and anonymous referees for invaluable comments.
MODALITY IS NOT EXPLAINABLE BY ESSENCE


University Nacional Autónoma de México (UNAM), México