

# **Essentialism Towards Necessity: A Critical Discussion**

**D i s s e r t a t i o n**

**zur**

**Erlangung des akademischen Grades**

**Doktor der Philosophie**

**in der Philosophischen Fakultät**

**der Eberhard Karls Universität Tübingen**

**vorgelegt von**

**Gaétan Bovey**

**aus**

**Romanel-sur-Lausanne, Schweiz**

**2024**

**Gedruckt mit Genehmigung der Philosophischen Fakultät  
der Eberhard Karls Universität Tübingen**

**Dekanin: Prof. Dr. Angelika Zirker**

**Hauptberichterstatter: Prof. Dr. Thomas Sattig**

**Mitberichterstatter: Prof. Dr. Fabrice Correia  
PD Dr. Tobias Wilsch**

**Tag der mündlichen Prüfung: 21.07.2023**

**Universitätsbibliothek Tübingen, TOBIAS-lib**

# **Essentialism Towards Necessity:**

## **A Critical Discussion**

**Gaétan Bovey**

A thesis presented for the degree of  
Doctor of Philosophy

Philosophisches Seminar  
Bursagasse 1  
D-72070 Tübingen

## Abstract

I start the present investigation by discussing the recent history of two notions that are at the center of contemporary metaphysics: *essence* and *metaphysical necessity*. According to an idea that was popular in the last half of the 20th century, *modalism*, essence is to be analyzed in terms of *de re* necessity. However, the viability of modalism seems to have seriously decreased as a consequence of Kit Fine's putting forth a strong challenge against it. On the one hand, Fine raised a series of substantial objections against modalism. On the other, he endorsed a whole new way to construe the relation between essence and necessity: he claims that necessity reduces to essence and not the other way around. According to the Finean essentialist view, then, essence plays the role of the *source* of—and therefore *explains*—necessity in the metaphysical structure of the world.

While Finean essentialism has received tremendous support by philosophers in the literature of the past 20 years, it is only recently that it has been the target of serious criticisms. One of the main goals of this dissertation is to discuss two central difficulties for Finean essentialism, and to see *whether* and *how* they can be addressed. To that effect, the first difficulty I introduce focuses on a rising concern according to which there is an *explanatory gap* between essence and necessity. Whereas this difficulty targets the *intelligibility* of Finean essentialism and whether it achieves its main explanatory task, the second difficulty pertains to the *extensional adequacy* of this theory; here, I argue that not *all* necessities can be accounted for in terms of essence. Given these problems, my next step consists in appealing to the notion of *generalized identity* in an attempt to find a construal of

essence that can overcome the difficulties I raise against Fine's theory. My conclusion, however, is that Finean essentialists are forced to work with a restricted version of their theory, and that a recent proposal made in the literature to address the explanatory gap is unsuccessful. I end the present investigation by expanding on the question of what the source of necessity is, and I present a framework where necessity can be generated from logical axioms and where essence transfers necessity.

Even though the conclusions I reach in my dissertation are, for the most part, negative towards Finean essentialism, I believe that they advance the debate in at least two different ways: on the one hand they challenge Finean essentialists to clarify and refine their claims and theory. On the other, they offer a novel and broader perspective on what kind of phenomenon necessity is and on the role that essence might play in the metaphysical structure of reality.

# Acknowledgments

For their continuous support and encouragements, I am grateful and indebted to my supervisors Thomas Sattig and Tobias Wilsch. As philosophers and human beings, they have continuously guided me on this (sometimes difficult) path, stimulated my mind with their ideas (and, for the most part, their objections and insightful remarks), and helped me overcome numerous obstacles that I thought were dead ends. I also want to specially thank them for their patience and understanding throughout all the ‘ups and downs’ of this beautiful thing that life is. To me, they have been like a compass in this journey and I could not have hoped for better mentors in a topic as intriguing as metaphysics.

I want to particularly thank the German Research Foundation (DFG) for funding my PhD research in Tübingen University within the project “Quellen der Notwendigkeit/Sources of Necessity” [Projektnummer 394341335] that I have been honored to be a part of.

This thesis would not have been possible without the help of many people. My heartfelt thanks to Fabrice Correia who has always been a source of inspiration throughout my entire studies in philosophy, and whose support has been a driving force all along; I am honored that he accepted to be part of the examination committee of this dissertation. Thanks to Alex Skiles for his precious time and advice in the final steps of the writing process, I would not have been able to make it without him. Being able to pursue doctoral studies would not have been possible

without what I have learned from Richard Glauser and Daniel Schulthess in Neuchâtel, and I thank them both for their support. Thanks to audiences of the Oberseminar in Theoretische Philosophie in Tübingen who have helped me improve my presentation skills and who have discussed some of the materials in this dissertation in a friendly and insightful manner: you made the experience in Germany feel like home. Thanks also to audiences in Neuchâtel for their comments on the third chapter and in particular to Olivier Massin for giving me the chance to present some of my ideas in his Colloque de Recherches. In no particular order, I also want to thank the following persons who have been a meaningful part of this experience with me at some point: Hong Yu Wong, Mañjuśrī, Vincent Grandjean, Hagen Braun, Daniel Gregory, Riccardo Baratella, Maria Scarpati, Michi Wallner, Nathan Wildman, Martin Glazier, and Cécile Rosat.

My deepest gratitude goes to my beloved ones. In particular to my wife, Yun, who has always supported me and loved me. Her beautiful mind and soul gave me all the strength and courage I needed during this journey. I also want to thank my mother, Isabelle, who has always believed in me and instilled in me values and passion for everything I have ever undertaken. Thanks to my father, Eric, for sharing my joys and fears during this meaningful journey. Finally, thanks to Bernard, Chantal, Gérald, and my Taiwanese family for their love and support.

# Contents

<b>Chapter 1: On essence, necessity, and how these notions are related</b>	<b>8</b>
1.I Quine on the unintelligibility of essentialism .....	11
1.II Kripke on reference, names, and modality .....	16
1.III A modal account of essence and Fine’s counterexamples .....	36
1.IV Finean essentialism .....	44
<b>Chapter 2: Finean essentialism and the explanatory gap challenge</b>	<b>50</b>
2.I Explanatory gaps in philosophy .....	51
2.II A gap between essence and necessity? .....	66
2.III On bridging the gap between essence and modality .....	78
<b>Chapter 3: On the necessity of essence</b>	<b>94</b>
3.I The necessity of essence .....	95
3.II Assessment of S1 .....	99
3.III Potential solutions on behalf of essentialists .....	107
3.IV Iterated essences without bearer identity .....	112
<b>Chapter 4: Essence and generalized identity</b>	<b>118</b>
4.I Generalized identity .....	118
4.II From essence to generalized identity .....	127
4.III The explanatory gap problem revisited .....	135
4.IV Correia and Skiles’s essentialist framework .....	144
<b>Chapter 5: Non-reductive essentialism revisited</b>	<b>156</b>
5.I Non-reductive essentialism vs. the gap: a reply to Wallner and Vaidya .....	158
5.II On the necessity of essence’s role in explanations of necessity .....	169
5.III Essence and transmissive explanations of necessity .....	177
<b>Bibliography</b>	<b>186</b>



# Declarations

I hereby declare that:

A) Portions of chapter 1 have been included in my MA dissertation “Essence, modality, and intrinsicality” that I submitted and successfully defended with honors in 2018 at the University of Neuchâtel (Switzerland).

B) Portions of chapter 3 have been published in the following journal:

Bovey, G. (2022). On the necessity of essence. *Philosophical Studies*, 179(7), 2167-2185. <https://doi.org/10.1007/s11098-021-01758-2>

In nil sapiendo vita iucundissima est.

Publilius Syrus

# Chapter 1: On essence, necessity, and how these notions are related

From the age of Antiquity in which Plato was writing until the present day, philosophers have always been interested in knowing the nature of the things that inhabit the universe, and even of what might lie beyond its possible boundaries. In no other domain of theoretical philosophy have such investigations taken a more fundamental and important place than in *metaphysics*. Indeed, metaphysical investigation has been especially interested in the notion of *identity* and with what is needed for a particular thing to be the very thing that it is. In both classical and contemporary metaphysics, most assume some form of *essentialism* when approaching these investigations. Indeed, these so-called ‘essentialists’ have primarily focused their research on the nature of things.

To put it roughly, essentialism is the thesis that things possess essential properties and/or essences. The essence of a given thing can be understood as the collection (or set) of metaphysically salient properties that make this thing *what it is*.<sup>1</sup> Thus, the essential properties of a given entity are special in the sense that they ‘characterize its nature’ and are ‘constitutive of its identity’—*i.e.*, essential properties provide the identity conditions of things. In contrast, the properties of a thing that are ‘merely incidental’ to a thing’s identity are often called its *accidental*

---

<sup>1</sup> As Yablo explains, the essence of a given entity is “an assortment of properties *in virtue of which* it is the entity in question”, and “a measure of *what is required* for it to be that thing” (Yablo 1987: 297).

properties. To summarize, it is enough at this point to bear in mind that, contrary to accidental properties, the essential properties of a thing are the properties that are part of the identity and the nature of that thing: they are central and of utmost importance for their bearers who *need* to possess them in order to be what they are.

But there is something special about the idea that things *need* their essential characteristics in order to be what they are. What does the term ‘need’ refer to? What kind of strength does it convey? Is it a matter of context or merely about language? In contemporary philosophy, it is generally accepted that a thing needs its essential traits as a matter of *absolute necessity*: it is impossible for  $x$  to fail to be  $F$  (where ‘ $F$ ’ is an essential trait of  $x$ ); anything that fails to be  $F$  is, as a matter of necessity, not  $x$ . Philosophers usually agree on the idea that this kind of necessity is metaphysical: there is no scenario in which a thing can cease to have its essential traits.<sup>2</sup> Some examples will help.

Consider the number 3: we can say quite a lot of things about that number, namely that it is the successor of 2, the predecessor of 4, and that it is an odd number. Now, would the number 3 still be the number 3 if it did not possess any of these characteristics? Intuitively speaking, the answer is ‘no’. As a matter of fact, it is hard to imagine that 3 could fail to be odd and still be what it is if it were not odd. In philosophers’ jargon, this means that number 3 is *essentially* odd, because this characteristic pertains to the core identity of what it is for something to be the

---

<sup>2</sup> It is usually accepted that necessity comes in different kinds of strength. Physical necessity is, for instance, what is necessary according to some law of nature. However, even if it is physically necessary that  $p$ , it is generally admitted that it is metaphysically possible that not- $p$ . See Kment (2017).

number 3. But not only that. Since the number 3 cannot fail to be odd, philosophers say that 3 is *necessarily* odd. To take another, and more controversial example, think of Socrates. Intuitively, Socrates could have refused to take Plato as a student, or he could have had one more hair on his head than he actually had at a certain time of his life. These kinds of change do not seem to affect *what* Socrates is.<sup>3</sup> Rather, they are what philosophers call ‘accidental traits’: whether something possesses or fails to possess an accidental trait does not affect its core nature and identity. But if, on the other hand, one imagines Socrates failing to be a human being, matters are somewhat different: it seems that we are no longer thinking of Socrates if we attribute to him the characteristic of being a deer, for instance. Just like with the number 3 being odd, we think of Socrates’ humanity as an *essential* trait of his: Socrates would no longer be Socrates if he failed to be human, and philosophers tend to agree that Socrates is *necessarily* human.

It seems clear that, with the examples just given, there is a tight link between essence and necessity. For whatever characteristic is essential to something also appears to be necessarily possessed by that very same thing. This is certainly not a coincidence, and contemporary philosophers have, subsequently, tried to spell out, explain, and clarify what this link consists in over the last decades. As it is commonplace in philosophy, such investigations came with their shares of sophisticated and brilliant ideas, but also of substantial difficulties and obstacles.

---

<sup>3</sup> Of course, he would be different in the sense that he would not be a teacher of Plato, and that he would be a man with more hair on his head, but his ‘core’ identity would intuitively not be affected.

In the following sections, I will give a brief overview of the history of essentialism, necessity, and the relation between the two, starting with W. O. Quine, who thought that the former notion was unintelligible; this is the business of section 1.I. Afterwards, we shall see in section 1.II that, thanks to the works of Kripke and important clarifications given in his lectures, essentialism can be vindicated. The next step will be to formulate a proper account of essence in terms of necessity, and I will explain how Fine argues against it in section 1.III. In section 1.IV, I will present Fine's positive theory of essence and single out his main thesis which will be the focus of the discussion in the remaining chapters.

## 1.I Quine on the unintelligibility of essentialism

Quine's philosophical investigations flourished at a time when contemporary modal logic—the logic of what is necessary, possible, contingent, and impossible—was in its early developments. To that effect, it is crucial to bear in mind that, as far as the philosophical debates that took place before the 1970s are concerned, the vast majority of analytic philosophers believed that, when applied to statements, the notions of 'apriority', 'analyticity', and 'necessity' were coextensive.<sup>4</sup> Briefly, a truth has all three features if it can be known independently from experience, if it is true solely in virtue of the meaning of the words that compose it, and necessary if it cannot fail to be true (*i.e.*, its falsity is impossible). This means that a truth like 'all bachelors are married' would have all three features, but not a truth like 'Socrates is

---

<sup>4</sup> Similarly for the parallel notions of 'aposteriority', 'syntheticity', and 'contingency'. But I will not discuss them in the present investigation.

human', since the meanings of the words alone and how they are combined does not suffice to make it true, and we cannot infer by purely *a priori* means that Socrates is human. Thus, according to this view, 'Socrates is human' cannot be necessary. In addition to this, Millianism (Mill's direct reference theory) was considered to be officially replaced by Russell's theory of reference: *descriptivism*. (The theory according to which the meaning of a proper name is given by a set of definite descriptions.<sup>5</sup>)

Next, Quine argued that modal operators had to be understood as metalinguistic predicates only, since he believed that a non-linguistic construal of modal notions would lead to a revival of Aristotelian essentialism, a doctrine he deemed *unreasonable* (Quine 1961: 156). According to Quine, essentialism is the doctrine according to which

[a]n object, of itself and by whatever name or none, must be seen as having some of its traits necessarily and others contingently, despite the fact that the latter traits follow just as analytically from some ways of specifying the object as the former traits do from other ways of specifying it. (*Ibid.*: 155)<sup>6</sup>

---

<sup>5</sup> This is a rough outline of what descriptivism was at that time. Since then, the theory has been refined and improved, but these matters will not substantially affect the present discussion and the arguments I am about to present.

<sup>6</sup> The way Quine characterizes 'Aristotelian Essentialism' is not faithful to what Aristotelian essentialism is really about. Marcus, for instance, considers Quine's characterization as 'presumptuous' (Marcus 1971: 187). For the present discussion, however, it will be of no harm to read Quine as simply characterizing 'essentialism' *tout court*, and to leave the discussion of whether it is faithful to Aristotle or not for another time. See Cohen (1978), Matthews (1990), and White (1972) for discussions on that matter.

(Note that the traits that ‘follow analytically from some ways of specifying an object’ are for instance whether a number satisfies open formulas like ‘ $x < 7$ ’ or ‘ $x > 7$ ’.) First, it is worth stressing here that with his characterization of essentialism, Quine attacks *any* theory that countenances the distinction between essential and accidental properties (Marcus 1971: 187). Second, the crucial point made by Quine is that he believes that what essentialism is really about is at odds with the way *language* is used to make essentialist claims. That is, an essentialist will typically claim that, *independently* of the way we refer to it, a given thing possesses some of its characteristics essentially, and some contingently; and this is precisely what is meant by a characterization of essential and accidental properties in terms *de re* modality: the modal notions of necessity and possibility are assigned to particular things (*res*) as opposed to sentences (*de dicto*). Before going any further in the discussion, it will be helpful to temporarily define essential properties as Quine understands them:

**(QEP):**  $F$  is an essential property of a given entity  $x =_{def}$  independently of *any* language in which  $x$  is referred to (if at all),  $x$  necessarily has  $F$ .

**(QAP):**  $F$  is an accidental property of a given entity  $x =_{def}$  independently of *any* language in which  $x$  is referred to (if at all),  $x$  contingently has  $F$ .

Now, in the second part of the above quote, Quine thinks that essentialists pay no attention at all to the fact that, according to him, essentialist claims are *dependent* on *how* objects are described and/or referred to. Quine’s idea is that if John is



described in a certain way, it follows that John is essentially *F*, while if John is described in another way, it follows that John is contingently *F*.<sup>7</sup> Hence, if Quine is right, the immediate consequence of his remarks is that there is simply no way to *objectively* draw a distinction between the accidental and essential properties of a given entity. Indeed, he thinks that the idea that an entity possesses *in itself* essential and accidental properties is just *baffling*: necessity *de re* is unintelligible to him. He gives the following example to express his bewilderment:

Perhaps I can evoke the appropriate sense of bewilderment as follows. Mathematicians may conceivably be said to be necessarily rational and not necessarily two-legged; and cyclists necessarily two-legged and not necessarily rational. But what of an individual who counts among his eccentricities both mathematics and cycling? Is this concrete individual necessarily rational and contingently two-legged or vice versa? Just insofar as we are talking referentially of the object, with no special bias toward a background grouping of mathematicians as against cyclists or vice versa, there is no semblance of sense in rating some of his attributes as necessary and others as contingent. Some of his attributes count as important and others as unimportant, yes; some as enduring and others as fleeting; but none as necessary or contingent. Curiously, a philosophical tradition does exist for just such a distinction between necessary and contingent attributes. It lives on in the terms ‘essence’ and ‘accident’, ‘internal relation’ and ‘external relation’. [...] But,

---

<sup>7</sup> I am here setting aside the issue of whether the relevant essential claims are analytic or not.

however venerable the distinction, it is surely indefensible. (Quine 1960: 199-200)

What Quine aims to show with his example is that it is nonsensical to think that a given entity, say Jones, who happens to be both a mathematician and a cyclist, is necessarily rational and contingently two-legged, or that he is necessarily two-legged and contingently rational. The main point is that if Jones is described *as* a cyclist, then Jones is necessarily two-legged and contingently rational; but if he is described *as* a mathematician, then he is necessarily rational and contingently two-legged. Thus, the necessary (or essential) and contingent (or accidental) properties of Jones seem to vary according to how Jones is described, referred to, and thought of.

In the present case, the properties *being rational* and *being two-legged* cannot satisfy **(QEP)** or **(QAP)**. Quine takes this example to reflect a general point according to which it is not possible to make sense of the idea that an object is in itself necessarily or accidentally so and so independently of how it is described. That is, to say that Jones is essentially *F* or accidentally *G* is not determined by the way Jones is *objectively*, but at least partially determined by the way Jones is described. Hence, according to Quine, by saying that Jones is necessarily rational, an essentialist finds himself guilty of privileging a certain way of characterizing Jones to make his point; for different reasons, another essentialist might want to say that Jones is essentially two-legged, thus privileging this characterization over other ones. Hence if, as Quine suggests, essentialism consists in favoring specific ways to characterize entities, the idea that these entities have some of their

properties essentially and others accidentally involves making *arbitrary* distinctions between the properties a thing has. Put differently, to think that a thing can be necessarily so and so is to be confused about the relevant domains of application of modality: language versus ‘the’ world—as if someone wanted to use a massive hydraulic drilling machine to make meticulous watchmaking operations.

So what emerges from this discussion so far is that there is a conflict between whether the accidental and essential properties of a thing are dependent or independent of language. Quine argues that how a thing possesses such properties depends solely on a speaker’s intentions or purposes, not on facts about objective reality, while what he refers to as ‘Aristotelian essentialism’ is a doctrine according to which our language has no bearing whatsoever on the essential and accidental traits of the said thing. As I shall explain in the next section, Kripke will dissolve this conflict and clarify what the philosophical commitments are with respect to intuitions about essentialism. Importantly, this brief discussion highlights that, even if essentialism is *prima facie* a plausible and intuitive theory, its *intelligibility* has been the subject of severe criticism.

## 1.II Kripke on reference, names, and modality

Kripke’s works are known for having fundamentally changed analytic philosophy and, *inter alia*, vindicated essentialism against Quine’s skepticism. The depth of Kripke’s thought would require an entire thesis on its own to be appreciated fully;

consequently, for reasons of space and straightforwardness, I will focus on the essential aspects of his arguments for the present purposes.

In his 1970s lectures, the first important step that Kripke takes is to make a *conceptual* distinction between the notions of apriority, necessity, and analyticity. He has not much to say about the latter, except that it is a concept exclusively concerned with semantic matters: analytic statements are true (or false) in virtue of their meaning (1972: 39).<sup>8</sup> Rather, Kripke is mainly focused on apriority and necessity. Regarding the former notion, he notes that it is a concept which belongs to *epistemology*: to say that a statement is *a priori* is to say that it can be *known* independently of any form of experience—*i.e.*, that reason alone can uncover its truth-value. To give an example, the sentence ‘every child has had a mother’ is *a priori* because it arguably does not require any form of experience to know that it is true: if you know what being a child is, you know that it implies having a mother—reasoning alone suffices to know that the sentence is true.

With respect to the notion of necessity, on the other hand, Kripke explains that it is a mistake to think—as it was almost unanimously thought at his time of writing—that it is nothing more than the notion of apriority (or ‘analyticity’ for that matter). He says:

We ask whether something might have been true, or might have been false. Well, if something is false, it’s obviously not necessarily true. If it

---

<sup>8</sup> As I have explained it in the previous section, this is a distinction that Quine did not make.

is true, might it have been otherwise? Is it possible that, in this respect, the world should have been different from the way it is? If the answer is ‘no’, then this fact about the world is a necessary one. If the answer is ‘yes’, then this fact about the world is a contingent one. This in and of itself has nothing to do with anyone’s knowledge of anything. It’s certainly a philosophical thesis, and not a matter of definitional equivalence, either that everything a priori is necessary or that everything necessary is a priori. Both concepts may be vague. [...] But at any rate they are dealing with two different domains, different areas, the epistemological and the metaphysical. (*Ibid.*: 36)

Here Kripke explains that he is exclusively concerned with necessity as understood in a metaphysical—or counterfactual—sense. So if the notions of necessity and apriority are to be identified with one another, then it is not a trivial matter but a *substantive* one. Yet Kripke thinks that any argument for their identification must go wrong somewhere, since he will argue that the two notions are not even coextensive. In fact, Kripke claims that if philosophers like Quine have doubted the intelligibility of *de re* modality, it is because they have not appreciated the fact that necessity, analyticity, and apriority are notions that *can* come apart. Of course, he does not say that no truth can be analytic, *a priori*, and necessary—what he denies is an absolute coextensivity of these notions. Thus, in order to demonstrate that it is possible to have a clear understanding of *de re* modality, Kripke will present arguments to single out this notion by showing that there are both *a posteriori necessary truths* and *a priori contingent truths*. But before expanding on these

points, it is necessary to introduce some of Kripke's terminology, and to present his ideas regarding reference.

The first important step to take at this point is to explain the notion of *possible worlds*. The basic idea is that for every way our world could be, there is a possible world which is that way. Kripke is not explicit regarding the ontological status of such worlds, but he certainly rejects Lewis's modal realism according to which they are concrete entities no less than the universe we inhabit (Lewis 1986). To avoid losing track of the main discussion, I will not engage in substantial discussion regarding the ontological import of expressions like 'there is a possible world...' or 'in every possible world...'. For our present concern, it will be sufficient to interpret what is meant by a 'possible world' as a story that could have been true. Hence, modal facts can be understood in terms of facts about possible worlds in the following way:

**Necessity:** a statement is necessarily true just if it is true at every possible world;

**Possibility:** a statement is possibly true just if it is true at some possible world;

**Contingency:** a statement is contingently true just if it is true in the actual world but false at some other possible world.

The main idea is that when the notion of possible worlds is invoked, we take the actual world and the entities that we can identify in it as a ground to elucidate modal concepts and counterfactual situations where we wonder if circumstances could have been different for these entities.

Now that the notion of possible worlds is introduced, we can turn to Kripke's philosophical insights regarding reference. In *Naming and Necessity*, Kripke acknowledges that descriptivism helps address a number of systematic problems and puzzles related to philosophy of language, such as those surrounding: empty names, co-referring names, and negative existentials. These represent some of the core reasons why descriptivism was almost unanimously endorsed by philosophers in the first half of the twentieth century. However, despite these considerable theoretical advantages, Kripke believes that descriptivism is fundamentally flawed, and that it should be rejected as a theory of meaning for names. Kripke famously argues that, as opposed to what descriptivists think, we cannot always substitute a proper name with a definite description *salva veritate*. He attacks the following core thesis of descriptivism:

**(D1)** Every proper name is synonymous with at least one definite description.

It should be noted that for Kripke, **(D1)** presupposes that anyone who understands a name knows which definite description it is synonymous with. That is, in order to properly understand a name, it is mandatory to know at least one description with which this name is synonymous.

Following Soames (2004), it is customary in analytic philosophy to distinguish between Kripke's epistemic, semantic, and modal arguments against **(D1)**. I will now only focus on the latter in order to properly introduce Kripke's notion of rigid designation, which will be central for discussing and refuting Quine.

Kripke's modal argument is so-called because it purports to show that descriptivism gets certain modal facts wrong. To see why this is so, let us take for granted that the name 'Plato' is synonymous with 'the best student of Socrates'.<sup>9</sup> If **(D1)** is true, it should be possible to replace one term for the other in each sentence they appear in *salva veritate*:

**(M1)** If exactly one person was the best student of Socrates, then the best student of Socrates was the best student of Socrates;

**(M2)** If exactly one person was the best student of Socrates, then Plato was the best student of Socrates.

These sentences are both true. However, Kripke notices that there is a modal difference between propositions **(M1)** and **(M2)**: the former is necessarily true while the latter is contingently true. Indeed, **(M1)** is necessarily true because in every possible world where exactly one person was the best student of Socrates, then the best student of Socrates was the best student of Socrates. However, we can imagine a world, *w*, where Plato never got involved with philosophy because he chose instead to devote his life to being a bard, while it is another individual,

---

<sup>9</sup> This is just a toy example for sake of illustration.



Cratylus for instance, who happens to be the best student of Socrates instead. In such circumstances, proposition **(M1)** will still be true in  $w$ , but proposition **(M2)** will be false. Kripke rightfully concludes that since propositions **(M1)** and **(M2)** differ in truth-conditions (*i.e.*, their modal status), they therefore do not possess the same meaning. Therefore descriptivism must be false—*i.e.*, if **(D1)** was true, then **(M2)** should be necessary true. We can now turn to *rigid designation*.

For Kripke, the difference in truth-value between **(M1)** and **(M2)** reflects a deeper difference between two different categories of terms. Kripke thinks that there are in fact excellent reasons to think that some terms *necessarily* denote their referents, while other terms *contingently* denote their referents. To this effect, Kripke introduces the notions of *rigid designation* and *accidental* (or non-rigid) *designation*. This is Kripke's positive theory.

Roughly, the idea is that rigid designators are terms that have the same referent across *all* possible worlds, while accidental designators are terms that do not meet this criterion.<sup>10</sup> According to Kripke, proper names are rigid designators and descriptions are typically accidental designators. He explains that this distinction rests upon intuitions we have about names (Kripke 1972: 48): we can pick any proper name with which a description is associated, and we will see that the name is rigid and that usually the description is not. In other words, one simply has to answer the following question in order to know whether or not a term is a rigid designator: 'could someone other than the actual A have been the A?' If the answer

---

<sup>10</sup> Fitch provides the following precision: "The basic idea of rigid designation is one where the expression in question has a fixed designation [...] when considering various counterfactual situations." (Fitch 2004: 36)

is 'yes', then 'A' is an accidental designator; if the answer is 'no', then 'A' is a rigid designator. To give an example, think of 'Judas' and 'the disciple who betrayed Jesus'. While someone other than the disciple who betrayed Jesus could have betrayed him, no other person than Judas could have been Judas. That is, Judas could have not betrayed Jesus, but he could not have existed without being Judas. While this example correctly reflects Kripke's intuitions, there are nonetheless three important clarifications to be made.

First, there are exceptions with respect to definite descriptions because some of them are essentially rigid. For instance, descriptions which designate mathematical objects, like 'the prime number between 127 and 137' and other very specific examples such as 'the pyramid of Khafre'. Second, to say that 'Judas' refers to Judas in other possible worlds is not to say that the inhabitants of such worlds necessarily use 'Judas' to refer to Judas. That is, we can imagine counterfactual situations in which it is '11' or even 'Aristotle' that are used to pick out Judas rather than 'Judas'. What matters is that when we imagine another possible world, we are supposed to ask ourselves *how* we would describe this possible world using our *own* language, and not how it would be described in whatever language (if any) is used by the inhabitants of that possible world. Finally, a more precise definition of rigid designators is needed. For although we can safely assume that numbers exist across all possible worlds if they exist in any, this is certainly not the case for other entities such as you, me, and ordinary objects. Thus, it should not be thought that a proper name like 'Socrates' refers to Socrates even in worlds in which he does not exist – which is controversial. This difficulty is not insurmountable, and to avoid confusion, we can use Fitch's definition of rigid designators to get around it:

A term *d* is a rigid designator of an object *x* if and only if *d* designates *x* at every possible world where *x* exists and does not designate anything other than *x* at any possible world. (Fitch 2004: 37)

Given these clarifications, we can better understand why there is a difference of truth-value between sentences (M1) and (M2) above: because it picks out whoever happens to uniquely satisfy ‘the best student of Socrates’ (Cratylus, Simmias, Cebes, etc.), it is surely a contingent matter who satisfies the description ‘the best student of Socrates’, so this description is an accidental designator. Thus, in imagining a possible world where the best student of Socrates is not Plato, one is not imagining a possible world where Plato is not Plato. Rather, one imagines a possible world where a different individual satisfies the description ‘the best student of Socrates’. On the other hand, ‘Plato’ is a rigid designator, because in invoking counterfactual situations involving Plato, one imagines various aspects under which Plato could have been different than what he actually is; and in such mental processes, one need not imagine a counterfactual situation about the best student of Socrates.

At this point, a descriptivist could respond to the modal argument as follows: Kripke may have shown that ‘Plato’ is not synonymous with ‘the best student of Socrates’, but he has not conclusively proven that ‘Plato’ is not synonymous with any definite description whatsoever. Indeed, if – as Kripke himself acknowledges – there are rigid descriptions, then the descriptivist could find such descriptions to provide the meaning of names. A straightforward solution is to use an actuality operator to rigidify definite descriptions, like ‘the *actual* best student of Socrates’.

But first of all, this solution is problematic because it seems *ad hoc*. Second, Soames (2005) correctly observes that if descriptivists endorse this move in response to Kripke, then it is unworkable. Indeed, by rigidifying definite descriptions with an actuality operator, the descriptivist seems committed to the claim that the two following sentences have the same truth-value:

(K1) Kripke believes that Plato wrote *The Republic*;

(K2) Kripke believes that the actual best student of Socrates wrote  
*The Republic*.

Now imagine a possible world,  $w$ , which differs from our world with respect to the sole fact that the Taj-Mahal is one millimeter taller. We can safely assume that in  $w$ , Kripke believes (K1). However, it is absurd to think that in  $w$ , Kripke possesses any belief about the world we live in, and so (K2) is false in  $w$ . Although this topic deserves a more extensive discussion, I shall leave the debate about whether or not descriptivism can be sustained against Kripke's objections to another time. Given that there is a difference between terms that necessarily denote their referents and terms that do not, we are now equipped to examine how Kripke's insights can help one to address Quine's philosophical challenges.

I have just introduced two important notions: possible worlds and rigid designation. To say that analyticity, apriority and necessity are conceptually different is one thing, but I will now focus on how Kripke uses the notions of possible worlds and rigid designation to show that necessity is not coextensive with apriority and analyticity—*i.e.*, if we think about the set of all analytic truths, all *a priori* truths,

and all metaphysical truths, their elements *differ*. If sound, these arguments will allow us to single out the notion of metaphysical necessity and we will be able to address Quine's skepticism towards the intelligibility of the accidental/essential properties distinction.

The main argument that Kripke provides involves a classic example with two celestial bodies. In Antiquity, people observed every morning a star in the sky and named it 'Hesperus', and they observed every evening a star as well and named it 'Phosphorus'. Back then, people thought that Hesperus and Phosphorus were two different (*i.e.*, distinct) celestial bodies, but it was discovered much later on that these two celestial bodies were, in fact, one and the same: the planet Venus. This leads us to some philosophical puzzlements. First, it is a logical truth that Hesperus = Hesperus, and therefore necessary. In fact, it is a truth that anyone equipped with the relevant cognitive abilities can figure out purely *a priori*.<sup>11</sup> (The same can be applied to any entity:  $a = a$ , Socrates = Socrates, and so on.) However, one cannot discover by purely *a priori* means that Hesperus = Phosphorus; this took empirical evidence and quite some precise observation about the sky with relatively advanced materials like a telescope. We can now present the following argument:

- (1) Hesperus = Phosphorus
- (2) Necessarily, Hesperus = Hesperus
- (3) Therefore, it is necessary that Hesperus = Phosphorus

---

<sup>11</sup> Note that this is true so long as the person has whatever it takes to acquire the ability to use the relevant terms properly (which might require at least some perceptual experience).

Premise (1) is a true statement of astronomy, premise (2) is a logical truth, and the conclusion follows naturally from these premises via the *principle of substitutivity of identicals*. This principle is fundamental for extensional languages, because it states that co-referential expressions can be substituted for one another in each sentence in which one appears *salva veritate*.<sup>12</sup> A skeptic like Quine, however, will reject this argument on the grounds that proper names are synonymous with some definite description(s). If this is so, then premise (1) should be rephrased as follows:

(1\*) The morning star = the evening star

Consequently, a descriptivist will point out that there are possible worlds in which the morning star is not identical to the evening star, and that (1\*) is merely contingent and *both* synthetic and *a posteriori*. After all, there is likely a world where the morning star is, say, Jupiter, and the evening star Venus.

With respect to premise (2), Quine will (rightfully) consider it as a true (*de dicto*) analytic statement that is equivalent to:

(2\*) Necessarily, the morning star = the morning star.

---

<sup>12</sup> Epistemological contexts can also be seen as referentially opaque. For instance, even though Bruce Wayne is identical to Batman, we cannot substitute these proper names *salva veritate* in a knowledge-based example like the following: if 'John believes that Bruce Wayne is rich' is true, we have no grounds to infer that 'John believes that Batman is rich'; a mere substitution of co-referential terms here does not preserve truth, because John might be completely ignorant of the fact that Bruce Wayne is Batman.

Whatever is referred to as the morning star in any world is identical to the morning star (in that world). Given (1\*) and (2\*), it is clear why Quine deems the conclusion to be unacceptable. Since he believes that all there is to necessity is strictly concerned with linguistic/analytical matters, he cannot accept the following claim:

(3\*)      Therefore, it is necessary that the morning star = the evening star.

Even though (3\*) follows from the premises (1\*) and (2\*), this conclusion is neither analytically true nor true *a priori*, and it directly contradicts (the true and contingent) claim that the morning star = the evening star. To that effect, Quine argues that a conclusion as absurd as (3\*) is a sufficient ground to see modal contexts as referentially *opaque*. That is, as opposed to referentially *transparent* contexts, an opaque context is a context in which the principle of substitutivity of identicals does not always hold. As we can see, if we hold a firm grip on the idea that necessity, analyticity, and apriority are coextensive notions, and that every proper name is synonymous with a definite description, the above argument can hardly make sense. For, even though (1\*) and (2\*) are both true, substituting ‘the morning star’ with ‘the evening star’ in modal contexts yields to sentences (*e.g.*, (3\*)) that are, under a Quinean understanding of these matters, not true.

However, with the notions of possible worlds and rigid designation in mind, we can see exactly what went wrong with this Quinean reconstruction of the argument. First, we have to keep in mind that necessity should be seen as having a purely

metaphysical interpretation, rather than being a cluster that embodies semantic and epistemic matters as well. Next, Kripke agrees that the truth of (1) can only be established *a posteriori*: we can hardly discover that both the morning and the evening stars—*i.e.*, Hesperus and Phosphorus—are the same celestial body without empirical evidence. Quite unsurprisingly, the key question for Kripke is whether (3) is true and follows from (1) and (2). And since Kripke’s metaphysical argument shows that descriptivism is not a suitable theory of meaning for proper names (*cf.* Kripke’s modal argument above), a counterfactual situation in which the sentence ‘the morning star is the evening star’ is false does not by itself entail the falsity of (3). Quite the contrary: thanks to rigid designation, Kripke argues that (1) is not only true in the actual world, but true in *every* possible world because ‘Hesperus’ and ‘Phosphorus’ are rigid designators that pick out the *same* celestial object, namely Venus.

Substituting ‘Hesperus’ and ‘Phosphorus’ for definite descriptions yields to sentences that are contingently true like (1\*). But asking whether the evening star could have not been identical to the morning star is different from asking whether Hesperus could have not been identical with Phosphorus. In the case of definite descriptions, they denote different objects across possible worlds, while Kripke explains that names denote the same thing across all possible worlds.<sup>13</sup> Thus, by letting go of descriptivism and using only rigid designators, we can see that the truth of (3) holds and, as such, (3) is *both* a necessary and *a posteriori* truth. Hence, contrary to what (most) philosophers at the time thought, Kripke demonstrates with

---

<sup>13</sup> I set aside special cases of rigid definite descriptions here.



this example that necessity is not coextensive with apriority, and that therefore a necessary truth is not *ipso facto* an *a priori* truth.

Given the above remarks, we can focus on how Kripke engages the debate about essentialism against Quine. He does so by drawing inspiration directly from Quine's mathematician-cyclist example. Kripke asks:

Was it necessary or contingent that Nixon won the election? [...] But this is a contingent property of Nixon only relative to our referring to him as 'Nixon'[...]. But if we designate Nixon as 'the man who won the election in 1968', then it will be a necessary truth, of course, that the man who won the election in 1968, won the election in 1968. Similarly, whether an object has the same property in all possible worlds depends not just on the object itself, but on how it is described. So it's argued. It is even suggested in the literature, that though a notion of necessity might have some sort of intuitions behind it [...], this notion [of a distinction between necessary and contingent properties] is just a doctrine made up by some bad philosopher, who (I guess) didn't realize that there are several ways of referring to the same thing.  
(Kripke 1972: 40-1)

Kripke explains that skeptics like Quine give no credit to the *intuitive* distinction between accidental and essential properties: in their eyes essentialism is just some kind of antiquated doctrine that should ultimately be cast aside in the light of serious philosophical considerations. In using different terms to pick out Nixon, it

seems as though one can affect what is true of him necessarily and what is true of him merely contingently. Kripke presents how an ordinary man and a skeptic would discuss the matter with each other:

Suppose that someone said, pointing at Nixon, ‘That’s the guy who might have lost’. Someone else says, ‘Oh no, if you describe him as “Nixon”, then he might have lost; but, of course, describing him as the winner, then it is not true that he might have lost’. Now which one is being the philosopher here, the unintuitive man? It seems to me obviously to be the second. The second man has a philosophical theory. The first man would say, and with great conviction, ‘Well, of course, the winner of the election might have been someone else. The actual winner, had the course of the campaign been different, might have been the loser, and someone else the winner [...]. So, such terms as “the winner” and “the loser” don’t designate the same objects in all possible worlds. On the other hand, the term “Nixon” is just a name of this man’. When you ask whether it is necessary or contingent that Nixon won the election, you are asking the intuitive question whether in some counterfactual situation, this man would in fact have lost the election. (*Ibid.*: 41)

Here, Kripke shows by means of an ordinary situation that our intuitions about essentialism are more than plausible and, most importantly, should come *first* when it comes to inquire about such questions. Indeed, there is nothing wrong or puzzling in imagining that Nixon might have lost the presidential elections of 1968,

and that Humphrey (or anyone else), for instance, might have won instead of him. But even if we happen to refer to Nixon with the description ‘the winner of the 1968 presidential election’, we still understand that we are referring to Nixon, who might have lost against his opponent. Similarly, when we use the description ‘the number of Swiss ministers’, we understand that we are referring to the number 7, which might not have numbered the Swiss ministers. Such counterfactual situations are obviously not unintelligible. In fact, as human beings we engage in counterfactual reasoning such as this on a regular basis, and our capacity to do so strikes us as hardly worth commenting on. We need not be familiar with any kind of philosophical theory: today it is sunny, but it might have been rainy; the number of planets is even, but it might have been odd, and examples proliferate.

For this reason Kripke goes on to argue that, as opposed to an ordinary person, it is the skeptic who defends a confused philosophical position:

If someone thinks that the notion of a necessary or contingent property (forget whether there are any nontrivial necessary properties [and consider] just the meaningfulness of the notion) is a philosopher’s notion with no intuitive content, he is wrong. Of course, some philosophers think that something’s having intuitive content is very inconclusive evidence in favour of it. I think it is very heavy evidence in favour of anything, myself. I don’t really know, in a way, what more conclusive evidence one can have about anything, ultimately speaking. But, in any event, people who think the notion of accidental [and essential] property unintuitive have intuition reversed. (*Ibid.*: 41-2)

Quine scrutinized the doctrine of essentialism through the lens of a well established philosophical view—descriptivism—together with the assumption that the notion of necessity just is the notion of analyticity and apriority.<sup>14</sup> It would be unfair to blame Quine for thinking and arguing that way. Descriptivism is indeed a powerful theory and I think that Quine was totally entitled to express several doubts about giving it up without substantial arguments. With respect to this point, although Kripke shows that descriptivism is not a suitable theory of meaning, he does not propose a solution for philosophical puzzles regarding empty names and negative existentials. Nonetheless, Quine clearly failed to recognize that in asking whether something is necessarily *F*, one does not want to know whether it is analytic that this thing is *F*. And, in fact, Kripke showed that in some cases, necessity is not coextensive with apriority (see the above argument with Hesperus and Phosphorus). Kripke pursues his argument against Quine and asks:

What's the difference between asking whether it's necessary that 9 is greater than 7 or whether it's necessary that the number of planets is greater than 7? Why does one show anything more about essence than the other? The answer to this question might be intuitively 'Well, look, the number of planets might have been different from what it in fact is. It doesn't make any sense, though, to say nine might have been different from what it in fact is'. (*Ibid.*: 48)

---

<sup>14</sup> Let me stress that some philosophers think that Quine was not fully committed to descriptivism, and that he would sometimes endorse the view for a specific purpose and abandon it for another one. Yet for the scope of the present work, I will not engage into this debate. I redirect the reader to Neale's 'On a Milestone of Empiricism' (2000) for a thorough discussion of these matters.

It is the notion of rigid designation that vindicates the intelligibility of essentialism. Indeed, a rigid designator picks the same entity across all possible worlds (where that entity exists). So to ask whether 9 has the property being odd essentially is to ask whether the sentence ‘necessarily, 9 is odd’ is true. However, if an accidental designator, *a*, is used to denote 9, then the truth-value of the sentence considered does not depend solely upon 9 anymore, but rather on which entities are picked out in which world by *a*; and this is simply irrelevant as far as the essential properties of 9 are considered. As Kaplan explains, rigid designation is “prior to the acceptance (or rejection) of essentialism, not tantamount to it” (1986: 265).

So as opposed to what Quine thought, as far as *de re* modality is concerned, it *does* make sense to ask whether an object is necessarily *F independently* of how it is specified. Indeed, just as the question of whether 9 is necessarily odd is not equivalent to the question of whether the number of planets is necessarily odd, the question of whether Nixon could have lost the 1968 elections is not equivalent to the question of whether the husband of Pat Ryan could have lost the 1968 elections. Quine’s specific example involving Jones, who happens to be both a mathematician and a cyclist, is also not bewildering at all: neither being a cyclist nor being a mathematician are essential properties of Jones, because it makes sense to imagine counterfactual scenarios in which Jones *did not* engage in one or both of these activities. On the other hand, there is something clearly wrong in imagining a scenario in which Jones is not Jones. By holding a firm grip on the *de re/de dicto* distinction applied to the notion of metaphysical necessity, there is nothing baffling about essentialism. Hence, we have good reasons to think that we can draw a distinction between accidental and essential properties with the help of modality.

We can now use Kripke's observations to prove a general and fundamental principle that governs identity:

**The necessity of identity**       $\forall x \forall y (x = y) \supset \Box(x = y)$

(*Read: for every x, for every y, if x is identical to y, then it is necessary that x is identical to y.*)

The proof goes as follows:

- I.       $\forall x \Box(x = x)$
- II.      $\forall x \forall y (x = y \supset (\Box x = x \supset \Box x = y))$
- III.     $\forall x \forall y (x = y) \supset \Box(x = y)$

Premise **(I)** is simply the idea that everything is necessarily identical to itself, which would be quite difficult to argue against. Premise **(II)**, on the other hand, is an instance of Leibniz's principle of the indiscernibility of identicals, which states that identical objects share all and the same properties. In this case, the relevant property is *being necessarily identical to x*. The conclusion, **(III)**, follows by predicate logic. Importantly, the necessity of identity is true if and only if the variables stand for rigid designators and '□' expresses metaphysical necessity.

The necessity of identity is of great importance for contemporary metaphysics in general, including the present inquiry and especially in chapter 4. Thanks to the background on reference and possible worlds presented by Kripke, we have a good grip on the idea that, contrary to what Quine and skeptics alike thought, it makes

sense to think of objects as having properties necessarily independently of how they are described. The property of being identical to  $x$  is one example, but we can now investigate further to see whether some other properties are instantiated necessarily by their bearers. I will now turn to a modal account of essence and explain how Kit Fine argues against it.

### 1.III A modal account of essence and Fine's counterexamples

Curiously, Kripke did not himself formulate a modal account of essence. In fact, he demonstrated how essentialism could be sustained once modality *de re* is vindicated with a proper use of the notion of rigid designation. He also argued further that the origins of a being are essential to it, and this is basically all that we can find in his lectures. In order to find explicit formulations of a modal account of essence, we have to look into the works of Plantinga (1978), Stalnaker (1979), and Salmon (1981) to name only a few. For the present discussion, I shall use the following characterization:

Existential modal account **(EMA)**:  $x$  is essentially  $F =_{def.}$  necessarily, if  $x$  exists, then  $x$  is  $F$ .

It is crucial to have the existential condition in **(EMA)**. For without it we would have the following characterization:

$x$  is essentially  $F =_{def.}$   $x$  is necessarily  $F$

Aside from God and numbers maybe, the majority of objects and entities are not necessary beings—*i.e.*, beings that exist in every possible world. So if Socrates, for instance, does not exist in every possible world, then there are worlds in which he does not instantiate *any* property whatsoever. Consequently, if we do not include an existential criterion, then we have to admit that if  $x$  is a contingent existent, then  $x$  does not possess any essential properties, or we have to allow for an entity to have a property in a world in which it does not exist. This undesirable consequence is avoided with the existential condition of **(EMA)**.

What matters for the present discussion is that a modal account of essence aims to define essence in terms of necessity. As discussed in the beginning of this chapter, there seems to be an intimate link between essence and necessity. As such, **(EMA)** seems to be a promising characterization of what it is for a property to be essential to one of its possessors. In the literature, this characterization of essential properties helped elucidate different essentialist theories, notably the following (with Socrates as an example):

**Kind essentialism:** Socrates is essentially human if and only if, necessarily, if Socrates exists, then Socrates is human;

**Origins essentialism:** Socrates essentially originated from a specific zygote if and only if, necessarily, if Socrates exists, then Socrates originated from *this* specific zygote.



For the present inquiry, I will not be addressing whether or not these claims are true, nor will I argue or defend specific essentialist principles. What matters so far is that there is quite obviously a relation between essence and necessity, and that those who defend a modalist view argue that necessity is a more *fundamental* notion than essence, and thus a notion that can be used to define what the essential properties of things are. In this case, facts about essence are nothing over and above facts about necessity.

Even though an account like **(EMA)** seems promising, it faces its own share of difficulties. This time, the problem is not rooted in skepticism about the intelligibility of essence and/or necessity, but rather in what kind of question we are after when we seek to know what is essential to some given entity. To that effect, Kit Fine (1994) argues that the once so-called ‘traditional’ way to look at the relation between essence and necessity, which defined the former notion in terms of the latter, is misguided, because it is *extensionally* inaccurate. To that effect, Fine *does not* argue against the intelligibility of modalism—as Quine did towards what he called ‘Aristotelian essentialism’. Rather, he argues that modalism wrongfully assigns some properties as essential to their bearers while intuitively, it should not.

Fine targets both simple characterizations of essence in terms of necessity as well as **(EMA)** with a series of powerful counterexamples. To understand what these counterexamples show, let me briefly explain what Leibniz’s view on essence is. We can qualify Leibniz as a defender of *maximal essentialism*—*i.e.*, the idea that all properties instantiated by any entity are essential to it. This means that the number of hairs you possess, the volume of snow on the Matterhorn, the number of

continents, and the fact that one sneezed on January 1st are all essential properties to their respective bearers. The problem with this view is that it *overgenerates* essential properties: we do not want to say that you having a blue shirt today is essential to you; after all, you could have picked a red or a green one. Such properties are, I think, rightfully characterized as accidental. By endorsing this extreme version of essentialism, it is fair to say that Leibniz's theory is (at the very least) extensionally wrong: the set of essential properties of any given entity is too broad on his view, since it includes properties that should, intuitively, be excluded. On Leibniz's view, essential properties are no longer 'special' but rather *trivial*: take any property of any thing you can think of, and that property is essential to it. That seems wrong.

As it happens, Fine shows that **(EMA)** is subject to the very same kind of criticism, and his counterexamples have been widely received and appreciated "to a degree that is unusual for philosophical arguments" (Roca-Royes 2011: 66). Though **(EMA)** is more fine-grained than a maximal essentialist's view, Fine argues that it also has the undesirable consequence of qualifying properties as essential when that should not be the case.

To properly understand Fine's objections, we need to first ask ourselves what we are doing when we inquire into the essence of a thing. Fine explains that essence-related questions are (metaphysically-loaded) '*what*'-questions. So when metaphysicians are interested in Socrates's essential properties, they are after a series of answers that pertain to the question '*What* is Socrates?', in the most fundamental sense about Socrates's identity/nature. Thus, an account like Leibniz's

that simply overgenerates essential properties will not, by these lights, be successful because it does not provide accurate and informative answers. We do not have a better understanding of *what* is Socrates if we are told that he had a stain on his tunic at a certain time or that he had an even number of hairs when he drank the hemlock infusion. On the other hand, even if the following claims are debatable, to say that Socrates had such and such origins and that he is human are the kind of answers that take us closer to his nature, to *what* he fundamentally is.

As Fine himself explains:

[T]he metaphysician is not interested in every property of the objects under consideration. In asking ‘What is a person?’ for example, he does not want to be told that every person has a deep desire to be loved, even if this is in fact the case. What then distinguishes the properties of interest to him? What is it about a property which makes it bear, in the metaphysically significant sense of the phrase, on what an object is? It is in answer to this question that appeal is naturally made to the concept of essence. For what appears to distinguish the intended properties is that they are essential to their bearer. (1994: 2)

Fine identifies what I will label the ‘essentialist *desideratum*’, which basically states that any proper account of essence and essential properties should provide informative answers to ‘*what*’-questions. Should an account of essential properties fail to meet this *desideratum*, it would be better to have a more fine-grained theory

of essential properties. This is precisely Fine's angle against the modal-existential account of essence.

Fine identifies four main counterexamples to **(EMA)** which basically show that some properties are essential to their bearer, when intuitively they should not, because they fail to meet the above *desideratum*. Fine first asks us to consider some properties that Socrates and his singleton set possess. According to modal set theory, a necessary and sufficient condition for the existence of a given set is that its members exist. Thus, as a matter of necessity, the singleton {Socrates} exists just if Socrates exists. Furthermore, it is necessary that if both {Socrates} and Socrates exist, then Socrates belongs to {Socrates}. Consequently: it is necessary that if {Socrates} exists, then Socrates belongs to {Socrates}, and that if Socrates exists, then Socrates belongs to {Socrates}. According to **(EMA)**, it follows that it is essential to Socrates to be a member of {Socrates}, and essential to {Socrates} to possess Socrates as a member. I take this counterexample to be the most important one. The difficulty is that the essential connection between Socrates and his singleton set seems to be asymmetric: it is essential to {Socrates} to contain its sole member, but it is not essential to Socrates to belong to his singleton.<sup>15</sup> The problem with **(EMA)** is that it cannot make the important distinction between these necessary properties of Socrates and {Socrates}: it deems both the property of 'being a member of {Socrates}' and 'having Socrates as a member' as essential to

---

<sup>15</sup> As Fine himself puts it: "can we not recognize a sense of nature, or of 'what an object is', according to which it lies in the nature of the singleton to have Socrates as a member even though it does not lie in the nature of Socrates to belong to the singleton?" (Fine 1994: 5).

their bearers. But only one of these is essential to its bearer. Therefore, **(EMA)** fails. As Fine explains:

Strange as the literature on personal identity may be, it has never been suggested that in order to understand the nature of a person one must know to which sets he belongs. There is nothing in the nature of a person, if I may put it this way, which demands that he belongs to this or that set or which demands, given that the person exists, that there even be any sets. (*Ibid.*: 5)

Next, Fine considers the relation of numerical distinctness which is a necessary relation:  $\forall x \forall y$ , if  $x \neq y$ , then it is impossible that  $x = y$ . So, if Socrates and, say, the Eiffel Tower are numerically distinct, it follows that it is impossible that Socrates = Eiffel Tower. Therefore, it is impossible that Socrates exists and is identical with the Eiffel Tower. Yet, according to the modal account of essence, it follows that it is essential to Socrates to be distinct from the Eiffel Tower. However, bearing the relational property of being distinct from the Eiffel Tower is a property that absolutely everything—with the exception of the Eiffel Tower—instantiates. There is nothing special, so to say, to the fact that Socrates is distinct from that artifact, and citing this fact about Socrates does not help us answer the question of what Socrates is.

The third counterexample concerns all necessary truths. Since it is necessary that Socrates exists, then  $\phi$  holds (where  $\phi$  is a necessary truth), **(EMA)** will have as a consequence that it is essential for Socrates to be such that  $2 + 2 = 4$ , and similarly

for any logical truth and mathematical truth whatsoever. We can regroup these problematic trivial properties as such: being such that  $\phi$  (where  $\phi$  is a necessary truth). As with the two counterexamples above, it seems wrong to say that in order to figure out what Socrates is, we need to have a good grip on mathematical and logical truths. These properties involve things that are different from Socrates (numbers, sets, etc.) and they are not part of Socrates's nature.

Finally, let us to turn to existence. This counterexample follows directly from the characterization of **(EMA)** and is therefore specific to it. It turns out that existence is an essential property of each entity, because in *every* world in which an entity exists, it exists. In other words, it is necessary that every entity exists in whichever world it exists in. (Let me stress here that due to Kant's observations on these matters, one might have good reasons to reject characterizing existence as a type of property, and for that reason one might dismiss this counterexample.)

Fine correctly observes that, as far as essentialism is concerned, none of these four problematic properties are of any help for better understanding the nature of Socrates—quite the opposite. Generally speaking, the problem is that **(EMA)** entails that everything that is necessary to something is *ipso facto* essential to it, and this can certainly not be true, as we have just seen. In response, we could think of adding conditions that essential properties have to satisfy. I will not survey all the literature here, but there have been quite a few important improved modal accounts that aim at being more fine-grained than **(EMA)** in a way that allows them to address Fine's counterexamples.<sup>16</sup> Briefly, my preferred solution consists in

---

<sup>16</sup> Correia (2007), Cowling (2013), Denby (2014), Wildman (2013) and Zalta (2006).

adding a condition of *intrinsicity* (based on grounding) to the modal account. To that effect, the account provided in Bovey (2021) helps modal essentialists to respect the asymmetric relation between Socrates and his singleton as well as Fine's other counterexamples.

Fine is of course aware that modalism can be improved, and he very briefly surveys a few solutions of his own. However, he ultimately says that no successful modal account of essence can be found (1994: 8), and he argues that we should abandon modalism altogether and rethink what essence and its relation to necessity are. Let me say at this point that this is a rather unwarranted move from Fine. The reason is that from the mere fact that (EMA) has undesirable consequences, it does not follow that *no* modal account is possible. But regardless of one's view on (improved) modal essentialism, Fine highlights a crucial point, namely that even if a modal account of essence were possible, what the ultimate source of necessity is remains open.<sup>17</sup> I now turn to the positive aspects of Fine's theory, where he claims that, with a different understanding of how the notions of essence and necessity are related, we can set the grounds for a theory of the source of necessity.

## 1.IV Finean essentialism

This section will be brief, because Fine does not provide any argument for the views I am about to present in his 'Essence and Modality' nor in any later works.

---

<sup>17</sup> This is an important matter because there is likely a reason *why* a proposition is true in every possible scenario. For a discussion on this point, see also Wilsch (2017).

Basically, his ideas are pretty clear: he does not deny that there is a connection between essence and necessity; rather he thinks that, as opposed to what modalists think, it is necessity that is to be analyzed in terms of essence. As he explains it:

Why has the modal criterion let us down so badly? What is it about the concept of necessity which makes it so inappropriate for understanding the concept of essence? Certainly, there is a connection between the two concepts. For any essentialist attribution will give rise to a necessary truth; if certain objects are essentially related then it is necessarily true that the objects are so related (or necessarily true given that the objects exist). However, the resulting necessary truth is not necessary simpliciter. For it is true in virtue of the identity of the objects in question; the necessity has its source in those objects which are the subject of the underlying essentialist claim. [...] Indeed, it seems to me that far 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, be they concepts or individuals or entities of some other kind, 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. (1994: 9)

Thus, having left modalism behind for the reasons mentioned in the previous section, Fine claims that essence is *prior* to—and therefore more fundamental than



—modality rather than *vice versa*. We are now in a position to identify Fine’s main thesis: *essence is the source of necessity*.<sup>18</sup> This claim will be at the center of the following investigation, and the discussion in the next chapters are entirely dedicated to it.

Before moving any further, there is a couple of points that need to be clarified and addressed. First, Fine explains that even if essence is not analyzable in terms of *de re* necessity, we can still elucidate this concept with the help of *real definitions*.<sup>19</sup> Roughly, this means that the essence of some *x* can be understood as the collection of propositions that belong to the real definition of *x*. Thus, Finean essentialists typically endorse the idea that the essence of a thing is its real definition. Second, Fine made it clear that the main purpose of a theory of essence is to explain what it is to correctly answer ‘what’-questions of the relevant sort. However, he seems to be uninterested in providing us with any clear criteria for how to single out essential properties. Thus, how are we supposed to figure out which properties are essential to any given thing? While this will not be a substantial matter for the present investigation, it is nonetheless something that must be kept on the radar for any theorist of essence that is sympathetic to Fine’s framework. To that effect, I ask the reader to keep in mind that all the essentialist claims used in the present work will not be argued for but simply used as examples.<sup>20</sup> Third, when Fine speaks of

---

<sup>18</sup> Unless explicitly stated, I will use the terms ‘essentialism/essentialist(s)’ and ‘Finean essentialism/essentialist(s)’ interchangeably in what follows.

<sup>19</sup> Since Aristotle, it is generally accepted in philosophy that, just as we can define words, we can define things as well. To give an example, the real definition of Socrates includes the proposition that he is human: it (partly) defines what Socrates is.

<sup>20</sup> If, for instance, you do not think that Socrates is essentially human—as Mackie (2006) argues—then simply replace that claim with one you are more comfortable with—*e.g.*, 2 is essentially even, or the empty set is essentially empty, etc.

‘each class of objects’ we have to understand this in a way that includes all *kinds* of objects—abstract and concrete—since metaphysical necessity includes truths about logic, mathematics, ordinary objects and so on. So for instance if it is necessary that 2 is even, this is so because it is essential to 2 to be even; if it is necessary that Socrates is human, this is so because Socrates is essentially human, and so on. The *collective* essence of different things can also be the source of some necessary truths: the fact that the Eiffel Tower and Socrates are necessarily distinct is explained by the collective essence of the Eiffel Tower and Socrates. Last, there are different ways to interpret Fine’s thesis, and this is an actual point of controversy. Some philosophers think that necessity *reduces* to essence in the sense that necessity is nothing over and above essence: that any fact about what is necessary *ultimately* boils down to a fact about the essence of some thing(s).<sup>21</sup> Others, including Fine himself in later works, argue that necessary truths are *grounded* in essence truths. Be that as it may, all philosophers sympathetic to Fine’s conception of the relation between essence and necessity agree on the general idea that essence is the *source* of necessity and that real definitions can help us theorize about essence. For the time being, we can formulate the sourcehood claim more precisely as follows:

**Source:** For every truth  $p$ , if  $p$  is necessary, then  $p$  is necessary because it is essential to some  $x$  that  $p$ .<sup>22</sup>

---

<sup>21</sup> See Correia (2012).

<sup>22</sup> For simplicity, I restrict the formulation of **Source** to cases involving a single entity (as opposed to multiple entities like the collective essence of  $x$  and  $y$ , etc.).

Throughout this entire investigation, I will flag when the debate between reductionism versus non-reductionism is relevant and whether **Source** is to be read as a reductionist claim or a non-reductionist one. What matters so far is that the general idea captured by **Source** is that essence explains *why*, for *any* given truth  $p$ ,  $p$  is necessary, and that all necessary truths can be explained in terms of essence.

Now that we have identified Fine's main thesis, we are entitled to ask ourselves a series of important questions, some of which will come to haunt us later in the final chapter. In particular, *why* take essence to be the source of necessity rather than anything else? Is it because the modal account first highlighted a very plausible relation between necessity and essence, in the sense that we can hardly think of a property being essential to a thing and not being necessarily possessed by the same thing? This is not, in my opinion, a sufficient ground to think that, among all that there is, essence is *the* prime candidate to be the source of necessity. Even more so if we acknowledge the clear possibility that necessity does not only admit of one source but of multiple ones. From the mere fact that a reduction of essence to necessity succumbs to a series of difficulties, it does not follow that a converse relation between these notions holds. As a matter of fact, Lewis makes no mention at all of essence in his *On the Plurality of Worlds* (1986), which is a serious contender for a theory of the source of necessity. My point is that even if the notion of essence is a coherent one, and even if there is something to be said in favor of the view that essence is the source of necessity, it is still an *open question* whether essence is the source of necessity given that there are other contenders.

Be that as it may, my main focus in this investigation is not on the source of necessity in *general*, but solely on Fine's idea that essence is *the* source of necessity, and on whether it achieves what it is supposed to within contemporary metaphysics—*i.e.*, providing explanations of each and every necessity in terms of essence. In what follows, will primarily focus on two difficulties that **Source** faces, and discuss whether they can be overcome with additional clarification and/or arguments. To see whether **Source** is sustainable, we have to focus on what the theory is supposed to achieve, how it aims to do so, and whether it succeeds in doing so. As I will argue, Fine's very own account of essence is subject to criticisms that the reader should be familiar with by now, namely issues that pertain to *intelligibility* (*cf.* discussion on Quine above), which I will discuss in chapter 2, and issues related to *extensional accuracy* (upon which Fine bases his arguments against modalism), which I will address in chapter 3. In chapter 4, I will consider a different approach to essence—that of generalized identity—and see whether it can successfully overcome the two difficulties identified in the chapters I just mentioned. In the last and final chapter, I will take a step back and discuss some authors that have recently tried to defend a specific kind of Finean essentialism, and give my own view on other viable options regarding the role that essence can play in a framework where logic generates necessity.

## Chapter 2: Finean essentialism and the explanatory gap challenge

In the previous chapter, I explained that Finean essentialism is the idea that essence *is* the source of metaphysical necessity: take any proposition  $p$  you like, if  $p$  is necessary, this is so because some entity is essentially such that  $p$ . Recently, some philosophers have raised a challenge regarding the *intelligibility* of this thesis. Casullo (2020), Leech (2020), Mackie (2020), Noonan (2018), and Romero (2019) argue that an *explanatory gap* exists between essence and necessity. Clarifying what an explanatory gap is will be the primary focus of section 2.1 of this chapter. I will explain what kinds of challenges explanatory gaps pose in the philosophy of mind, metaethics, and the philosophy of action. This will help us understand what the explanatory gap between essence and necessity amounts to, as well as help us distinguish between epistemic and metaphysical gaps. In section 2.II, I will detail what the explanatory gap is in the essentialist context, and identify what kinds of challenges it faces compared to other gaps. Finally, in section 2.III I will evaluate how two different construals of Finean essentialism—reductionism and non-reductionism—can meet the explanatory gap challenge. I will also explain that the question of *what* essence *is* plays a substantial role if one takes the gap problem seriously and wants to bridge it.

## 2.I Explanatory gaps in philosophy

Explanatory gaps are found in several areas of philosophical theorizing, and it is sometimes difficult to pinpoint exactly what kind of challenge they pose in a given context and/or for a given theory. To begin with, the term ‘explanatory gap’ was coined by philosopher Joseph Levine in his 1983 article ‘Materialism and Qualia: The Explanatory Gap’. In that paper, Levine construes the explanatory gap as an *epistemic* difficulty for materialist theories of the mind, but, as I will explain it, some explanatory gaps also admit of a *metaphysical* reading. Importantly, the idea that there is an explanatory gap between essence and modality is fairly new in the literature and, consequently, it has not received as much attention as in other domains of philosophy. Thus, in order to identify what kind of difficulties skeptical philosophers are raising against Finean essentialist theories of modality, I will discuss what the explanatory gap challenge consists in for materialist theories of the mind as well as in normative contexts. This will prove useful in identifying the key characteristics of explanatory gaps and what kinds of strategies can be used to bridge them.

Following the works of Kripke on identity and necessity, Levine (1983) raises a difficulty for materialist theories of the mind. In sum, proponents of these theories argue that the mental is nothing over and above the physical: mental states are *identical* to physical states. To give a classical example, a materialist will typically endorse the claim that pain *just is* (or is ‘nothing other and above’) the firing of C-fibers in a sentient being. The latter fact is supposed to *fully* explain what pain *is*: if you are experiencing pain, this is because you are in a (neuro-physical) state of C-

fibers firing. Levine, however, argues that even though this might provide a good explanation of, say, the *causal role* of pain (that is, what, physiologically speaking, pain does *in* and *to* a particular sentient being), something crucial about pain is left unexplained by materialism, namely: pain's *qualitative* role—*i.e.*, “*why pain should feel the way it does!*” (*ibid.*: 357). Levine's point is that by identifying pain with C-fibers firing, the phenomenological aspects of pain are left out of the explanation, and the same is presumably true of other mental states such as pleasures, desires, joys, fears, and so on.

So, the explanatory gap challenge raised by Levine against materialists is the following: no matter what physical explanation of a mental state one attempts to give, it will always be insufficient to explain *why* and *how* it feels like to be in that state—*i.e.*, *why* and *how* a particular physical feature gives rise to a particular phenomenological feature. Pure physicalist explanations of the mental seem to always leave something unexplained about the mental, its ‘what-is-it-like-ness’. Levine's point, then, is that even if a particular mental state is identical to a particular physical state, the identity-theory does not and—according to Levine—*cannot* provide an explanation of the phenomenological aspects of the mental state. The problem, then, is that the connection between the physical and the mental is an obscure one, and that we do not know how physical events give rise to mental events.

Roughly, this is what the explanatory gap challenge amounts to in philosophy of mind. Before going any further, two key aspects of this gap have to be flagged for the sake of our discussion regarding essentialism. First, the gap seems to be

grounded in characteristics that pertain solely to physical states and solely to mental states, respectively. To illustrate, the former are *objectively* observable: when scientists study the brain states of a given subject, they can *empirically* gather information about the firing of C-fibers. For instance, one can measure C-fibers's firing *rate* (cf. Koltzenburg & Handwerker, 1994). However, such observations are not available when it comes to mental states such as *experiencing* pain, joy, and other mental states. As Nagel explains: "It is difficult to understand what could be meant by the *objective* character of an experience, apart from the particular point of view from which its subject apprehends it" (1974: 443). In fact, thinking of the 'objective feature' of an experience seems to be incompatible with what an experience *is* in the first place, namely: an event that is essentially *subjective*. Thus, the 'mark' of mental states is that they are—as opposed to (neuro)physical states—*subjective* in nature, while the 'mark' of physical states (partly) consists in the possibility of having external empirical access to them. In my opinion, this discrepancy renders the explanatory gap between the physical and the mental quite *vivid*: how could something being (at face value) intrinsically subjective be identified with a phenomenon that is publicly (*i.e.*, objectively) observable? *Prima facie*, there is no clear path that takes us from subjectivity to objectivity and *vice versa*: the *connection* (granted that there is one) remains blurry.

The second aspect is that there is an ongoing debate on what kind of conclusion can—and even *should*—be drawn from the explanatory gap in the philosophy of mind. In particular, this explanatory gap is, as Levine described it, first and foremost an *epistemic* difficulty. But some philosophers like David Chalmers argue that there being an epistemic gap between the mental and the physical is sufficient for there



being a *metaphysical* gap between these phenomena as well. In what follows, I want to clearly distinguish between the epistemic and the metaphysical aspects of a given gap, as well as identify what kind of explanatory demand they amount to.

Levine's epistemic gap primarily targets our capacity to give a conceptual analysis of the mental in terms of the physical, and he basically points out that it is likely that no satisfactory answer can (ever) be given. Block and Stalnaker also explain that the epistemic explanatory gap "is a gap in our understanding of *how* the physical facts make the mental facts true, a gap that would not be closed even if we accepted the thesis that the mental facts *are* made true by the physical facts" (1999: 3).<sup>23</sup> The problem, then, is that even if we *identify* a mental state with a particular physical state, it is still an *open question* why the mental state feels the way it does, and why *that*—and not *this*—mental state arises whenever the corresponding physical state occurs. This takes us to the crucial point regarding this explanatory gap: this 'lack in our understanding' is reflected in the fact that we cannot infer *a priori* a mental state from a physical state.<sup>24</sup> As Papineau explains, "[s]crutinize the physical facts about the behaviour of C-fibres as much as you like, and they will not tell you that it is painful to have your C-fibres firing" (2011: 7).

To this day, no satisfying explanation has been given in the literature, and the epistemic explanatory gap remains 'unbridged': it is what Chalmers coined 'The Hard Problem of Consciousness' (2007). The explanatory demand regarding this

---

<sup>23</sup> In this scenario, skeptics ask how and why does a physical state cause a particular mental state as well as reasons to know which "psycho-physical identity statements are true" (Levine, 1983: 354).

<sup>24</sup> Chalmers and Jackson (2001) both argue that materialism can be sustained only if there is a successful *a priori* conceptual analysis of the mental in terms of the physical.

epistemic gap is therefore the following: what are the reasons to *believe* that a particular physical state gives rise to that particular mental state? Put differently: granted that a physical event occurs, how do we know which mental characteristics it possesses? So far, it seems that nothing in our cognitive abilities allow us to infer *a priori* any kind of mental features from purely physical descriptions.

Thus, one possible conclusion to draw from the epistemic gap is that there being a gap between the mental and the physical is just a sign of our cognitive limitations. In fact, materialists argue that even though there is an epistemic gap between the mental and the physical, this has a no particular significance for how reality is structured, metaphysically speaking. As Kripke explains in his *Naming and Necessity*, we should distinguish between what is *a priori* and what is (metaphysically) necessary.<sup>25</sup> So, the fact that we cannot infer *a priori* that pain finds its source in C-fibers firing does not entail, in and of itself, that pain and C-fibers firing are, for instance, (metaphysically) distinct kinds of states. By comparison, just because we cannot infer *a priori* that water is composed of H<sub>2</sub>O, it does not follow that they are not one and the same thing.<sup>26</sup> If materialists are right, this means that even though our faculties are presumably not sufficiently developed to decipher the link between physical states and mental states, reality might well only consist of physical facts, states, properties and so on: our inability to know what the interaction between the mental and the physical consists in should not drive us to conclude that reality admits of two (metaphysically) distinct realms. If this is indeed where the story ends, the explanatory gap between the

---

<sup>25</sup> See chapter 1 for a presentation of Kripke's argument.

<sup>26</sup> I know this is a controversial point for some philosophers, but I merely use this as an example *qua* example.

physical and the mental is merely a strange ‘curiosity’ that we cannot, at least for now, comprehend as human beings; the difficulty is therefore merely conceptual and does not cross the ‘reality’ boundary.

Other philosophers like Jackson and Chalmers, however, argue that there being an epistemic gap between the mental and the physical is sufficient to conclude that there is a *metaphysical* gap between these notions as well. In this case, the consequence is that the mental and the physical are what Descartes considers to be two separate *substances* that cannot be identified with one another, because they belong to two different ontological categories. Thus, the challenge for materialists consists in providing an explanation of *how* the physical gives *rise* to the mental—of *how* both realms are connected/correlated. Such an explanation should tell a ‘story’ regarding the process(es) by which the physical realm operates to generate the mental realm. So if the epistemic gap is a serious matter, as the aforementioned philosophers think, then the materialist’s identity theory must be rejected, because there is a metaphysical gap that jeopardizes any chance for reduction to succeed. So if there is indeed a metaphysical gap, materialists have to find a way to bridge it with, for instance, psychophysical laws of nature, causal explanations, or even a *relation* such as (unproblematic) grounding.

The lessons to be learned from this brief discussion of the explanatory gap in the context of philosophy of mind are the following. First, there is something intrinsically special about the mental—namely, its phenomenological and subjective aspects—that distinguishes it from the physical. And that difference makes the connection between the mental and the physical somewhat obscure to

us: our cognitive abilities can hardly process *how* a particular physical state gives rise to a particular mental state (or *vice versa*), nor why mental states feel the way they do. Second, from these observations it is clear that the explanatory gap is first and foremost an epistemic difficulty that is best characterized by a lack of *a priori* entailment. The question of whether this entails that there is a metaphysical gap between the mental and the physical requires an investigation on its own and requires that materialists find a way (*e.g.*, via grounding, psychophysical laws, causal explanation, etc.) to bridge the gap. Leaving a metaphysical gap open amounts to leaving the door wide open to the skeptic's criticisms.

Let us now set aside the explanatory gap challenge for materialists, and turn to metaethical theories which also face the threat of an explanatory gap. As our toy example, consider the property of *goodness*. Clearly, goodness is a normative property since it consists in a specific evaluative status of particular human actions, attitudes, and traits when someone says that such and such action, mental state, or attitude is *good*. But some of the questions that lie at the heart of metaethical investigations are 'what is the nature of goodness?' or 'can goodness be explained?'. More precisely, is there a way to spell out what goodness is in entirely non-normative terms? Throughout history, many philosophers have been invested in these questions and have come up with sophisticated ethical theories to answer them. For instance, Epicurus argued that with abstract moral principles and theorizing only, one cannot comprehend what is (fundamentally) good or bad. Rather, he argued that *pleasure* consists in the greatest good. So in seeking and achieving one's pleasure, one does what is fundamentally good for oneself. According to Epicureans, then, we can see that they explained what goodness is in

non-normative terms—*i.e.*, in terms that *describe* what is the most pleasurable for a given subject. This theory provides a *reduction* of what goodness is in terms of what is pleasurable: for  $x$  to be good (to  $y$ ) is for  $x$  to be pleasurable (to  $y$ ).

Other metaethical theories might pursue a similar kind of reductive project, such as, for instance, identifying goodness with what satisfies one's innermost desires, what benefits the greatest amount of people, or what conforms with Kant's categorical imperative. What these theories all have in common is that their aim consists in providing an account of goodness (and badness, for that matter) in non-normative terms. But this is problematic.

Moore (2004) explained that all reductive theories about goodness are inevitably doomed to fail to capture what goodness is, because no matter what goodness is reduced to, it remains an *open question* whether the target phenomenon is good or not. For instance, if you say that goodness just is pleasure, then, according to Moore, it remains an open question whether pleasure is good (*tout court* or for some individual). Importantly, asking such a question is not a symptom of a failure to understand what the reductive thesis tries to establish or a misinterpretation of what the relevant concepts involved in the theory considered actually *mean*. Rather, Moore's point is that someone who asks whether pleasure also happens to be good is asking a *genuine* and *intelligible* question, as if someone were wondering 'I can see that this is a pleasurable action, but is it also a good action?'. Thus, the problem is that thinking that  $x$  is pleasant is not the same as thinking that  $x$  is good. One can fully grasp that something is pleasant and yet fail to assess whether it is good or bad: surely, procrastinating is pleasant but probably not good in every

circumstance. Just like with the gap in philosophy of mind, the difficulty here is also first and foremost epistemic: when Moore explains that it is an open question whether something pleasant also happens to be good, what is meant is that it is *a priori* open that something pleasant is good. Put differently, one cannot infer, simply with a sufficient grasp of the relevant concepts involved and by *a priori* means, that pleasure (or whichever purely descriptive concept one reduces goodness to) is good (or bad).

The same sort of puzzlement about what kind of conclusion should be drawn from the epistemic gap in philosophy of mind arises in the metaethical context, namely: is the lack of *a priori* entailment from the descriptive to the normative merely a symptom of cognitive limitations and knowledge of the relevant concepts, *or* is it a symptom of a more fundamental problem that finds its source in metaphysical reality? If the former, then we do not only need reasons to believe that normative facts just are descriptive facts, but we also need to know which identity claims are true. If the latter, then the purely metaphysical explanatory demand partly consists in providing a theory of normative properties, and to explain why, granted that a certain description of reality holds across a set of worlds, the same set of normative constraints hold as well. In other words, *why* and *how* a particular description or portion of reality is always assigned with the same normative property.

So if a metaphysical gap is at stake, this means that the properties of goodness/badness cannot be reduced to non-normative properties. In this case, the challenge for proponents of the reductivist theory is to close the gap with the help of, for instance, a relation like grounding, supervenience, or, in this particular context,

perhaps with some kind of moral law. This is the normative explanatory gap in a nutshell: identifying the property of goodness with a non-normative property leaves something unexplained, just as with the phenomenological aspects of mental states that are identified with physical states (of the brain). To say that  $x$  is pleasurable is not equivalent to saying that  $x$  is good, and therefore such reductive theories fail to give a proper account of what goodness consists in so long as the gap remains unbridged.

Before going any further, there is a related problem that I want to briefly discuss as I think that doing so will help us identify the problem essentialists face even more vividly. Hume famously argued that one cannot derive an 'ought' from an 'is'. The idea is that from a given set of facts such as 'Mary drinks fifteen cans of Dr. Pepper every week' one cannot infer that 'Mary ought not to drink fifteen cans of Dr. Pepper every week'. I think it can safely be assumed that the normative claim has some truth to it: for health reasons, Mary *should not* have such habits. But the problem is that it seems that one can rightfully infer the 'ought'-statement only if one (tacitly) presupposes an evaluative/normative claim—*e.g.*, that large quantities of sugar are bad for a human being's health. From the mere fact that Mary drinks fifteen cans a week, the evaluative claim does not necessarily follow. Furthermore, the very same evaluative/normative presupposition that is assumed as an additional claim is not itself derived from non-evaluative/normative facts. As Sayre-McCord explains "[i]f Hume is right, every valid argument for an evaluative conclusion either includes or presupposes some evaluative premise. And, as a result, there is no value-neutral argument for an evaluative conclusion" (2011). As discussed with philosophy of mind and normative theories, the problem just highlighted shows

that some special kinds of features cannot be straightforwardly derived from claims that make no mention of the said features. More precisely, the difficulty is that there is a lack of *a priori* entailment from the ‘is’-fact to the ‘ought’-fact—something that does not occur *if* the former fact is ‘tacitly’ wearing a normative garment. As we will see below, this is one of the main difficulties that skeptics raise against essentialists who wish to explain modality in terms of essence.

So, generally speaking, when facing an explanatory gap between two phenomena, there is typically (at least) one characteristic feature of the target phenomenon being reduced that poses a difficulty—*e.g.*, it resists a conceptual analysis, it cannot be explained in different terms without leaving something important out, it fails to be accounted for in the target language, and so on. Moreover, we have seen that explanatory gaps are first and foremost epistemic in nature and that this is characterized by a lack of *a priori* entailment between the notions involved in a specific reductive framework. What kind of conclusion should be drawn from an epistemic gap is a difficult issue, but I would like to briefly say something about this now that some of the key characteristics that pertain to explanatory gaps have been identified.

The first point is that (some) epistemic gaps are somehow ‘cheap’ and can be found (almost) everywhere. For instance, Schaffer (*forth.*) argues that even in the case of water and H<sub>2</sub>O, we have such a gap. His rationale is that we can analyze the concepts of hydrogen and oxygen as much as we like, nothing warrants us in concluding that a given chemical H<sub>2</sub>O compound is water. This might sound like an extreme philosophical view but it is not far-fetched. In the literature, the gap



challenge was first construed at the level of the (alleged) connection between psychology and the so-called ‘hard’-sciences—with the wish to have a uniform theory where everything ultimately boils down to what belongs to the physical (or ‘physical events’ if you prefer). Given that there is no straightforward reduction of mental events to physical ones, *why* should it be in any way less obscure to reduce facts about chemistry, biology, and so on to purely physical facts regarding mere particles arranged in a certain way? As soon as we start to give a proper look at different theories, we can see that epistemic gaps are (almost) everywhere.

However, if we acknowledge the fact that epistemic gaps arise more frequently than we might initially think, then there is an important difference between Schaffer’s H<sub>2</sub>O case and the gaps I have described above. In particular, one has the sense that, in the case of pain, there is no way to further describe this mental state unless we appeal to more terms from the phenomenology handbook. With mental states, it seems that we can never get out of subjectivity: the kind of description we start with seems to entrap us and stick with us, and there does not seem to be a way out. Similar reasoning applies to the normative and ‘is-ought’ gaps: we seem to never be able to get rid of normative terms when we try to describe goodness or what one ought to/not to do. With the H<sub>2</sub>O case, however, things are substantially different: we are not entangled with some characteristics that are indescribable in different terms, for we stay in the same ‘realm’; what is obscure is the link between water and H<sub>2</sub>O, and in this case, an explanation is possible. As Schaffer explains, we can bridge such gaps with the help of *grounding*: even though we cannot infer *a priori* that water is H<sub>2</sub>O, we can clarify the connection between water-facts and

H<sub>2</sub>O-facts via the relation of ground; nothing seems to be left unexplained in this kind of scenario.<sup>27</sup>

The second point is that if an epistemic ‘flavor’ is all that there is to explanatory gaps, then this makes them relatively uninteresting. After all, we can just load this difficulty on the back of our cognitive abilities and simply make a bet on the future of technology and research development to understand what we have been missing so far, period. But it seems implausible that the explanatory gaps I have described above merely stop at the conceptual level, and I think that there is something that feels not quite right in saying so. Some gaps are probably merely epistemic, as with Schaffer’s H<sub>2</sub>O case and in this case reduction/identification is possible. But when materialists say that pain *just is* the firing of C-fibers, they are clearly leaving something out: the phenomenological aspects of pain. And these aspects are not merely conceptual, they vividly occur in reality within the boundaries of a sentient being’s body. An even more puzzling question, for instance, is why some thoughts can trigger a state of sorrow in someone: can this be captured in purely physical terms? When Mary goes out for the first time from her black and white room—a place where she stayed her entire life up to that point—is she really not experiencing something entirely new when she sees red for the first time, even though she knows everything she has to know—physically speaking—about redness and other colors? (Jackson 1982) Hence, I think that such explanatory gaps should be taken very seriously: after all, they are *gaps*, and we cannot simply stipulate a reductive claim (*e.g.*, with the help of the identity relation) between two phenomena (that we characterize with different concepts) without providing a more

---

<sup>27</sup> For simplicity, I set aside scepticism about grounding here.

substantial story about the underlying metaphysical mechanisms that warrant the existence of such a reduction. In fact, what I consider to be ‘serious’ gaps make reduction/identification impossible. So even though gaps can be found almost everywhere, we have to be careful because some of them are rooted way deeper than the conceptual level.

Thus, explanatory gaps raise questions at (at least) two different levels. On the one hand, there are questions that pertain to epistemology and which revolve around reasons to believe that two phenomena are related in an explanatory way or that one phenomenon is nothing over and above the other. In this case, even though reality might just consist of only *one* of these two phenomena, there is an incapacity for humans to obtain knowledge of that fact, to know that identity claims about them are *true*. On the other hand, there are metaphysical questions that are centered around the underlying mechanisms/processes by which a given phenomenon operates when giving rise to another phenomenon. This includes epistemological puzzlement, but metaphysical gaps go way beyond it since they concern how reality itself is structured.

It can be difficult to properly separate the metaphysical matters from the epistemic ones, and I think that we should keep in mind three questions when facing an explanatory gap between two phenomena  $\phi$  and  $\psi$ :

**Substantive question:** What are the characteristic traits of  $\phi$  that prevents a straightforward explanation/reduction of it in terms of  $\psi$ ?

(*E.g.*, the mental's phenomenological/subjective trait that resists explanation in physical terms.)

**Epistemic gap question:** What *reasons* are there to *believe* that one of  $\phi$  or  $\psi$  can be reduced to the other? (*E.g.*, that every mental state can be explained in physical terms, that every normative fact finds its source in a descriptive, naturalistic fact, or that  $\phi$ s just are  $\psi$ s.) In sum, the question here is: '*how* do we know that  $\phi$  and  $\psi$  are related?'

**Metaphysical gap question:** What are the underlying mechanisms/processes by which phenomenon  $\psi$  generates/gives rise to  $\phi$ ? (*E.g.*, how and why does a physical state give rise to that particular mental state?) As opposed to the epistemic question, the metaphysical question consists in figuring out *whether*  $\phi$  and  $\psi$  are indeed related.

The first important task consists in answering to the substantive question, and then to find out whether we are facing a mere epistemic matter or a metaphysical one. As I have explained, epistemic gaps are 'cheap' when it comes to any theory that aims to reduce one phenomenon with another: they can be found nearly everywhere. However, when it comes to metaphysical gaps, we need to find one (or more) principle to bridge them—via grounding, supervenience, laws, etc. In such cases, reduction or identification are no longer available moves. Appealing to these when facing a metaphysical gap leaves no room for explanation: it is the same as saying to someone 'you're wrong' without explaining why. I now turn to the

explanatory gap that essentialists face, and I will focus exclusively on the metaphysical gap between essence and modality.

## 2.II A gap between essence and necessity?

Recently, a series of criticisms targeting the intelligibility of the sourcehood thesis have been presented in the literature, and specifically against the idea that there is a suitable notion of real definition that can fulfill the role of the source of necessity. Some philosophers express skeptical concerns regarding *whether* and *how* essence generates necessity. As we shall see, the explanatory gap worries that threaten some theories in philosophy of mind and metaethics surface under a different—and yet familiar—guise for essentialists.

For the sake of the discussion, I will quote Fine as our starting point:

I shall [...] argue that the traditional assimilation of essence to definition is better suited to the task of explaining what essence is [than is the modal account of essence]. It may not provide us with an analysis of the concept, but it does provide us with a good model of how the concept works. Thus my overall position is the reverse of the usual one. It sees real definition rather than *de re* modality as central to our understanding of the concept. (1994: 3)

And he says further:

it seems to me that far 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, be they concepts or individuals or entities of some other kind, 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. (*Ibid.*: 8)

The first quote details how Fine understands essence, and the second quote consists in Fine's account of metaphysical necessity in terms of essence. From these passages, it follows that any necessary truth follows from the essence (understood here as a real definition) of some thing(s). The following principle will help us navigate through the criticisms that will be discussed shortly:

**(NP)** If being (an) *F* is an essential property of *x*, then being (an) *F* is a necessary property of *x*.

Importantly, all the philosophers that I will discuss do not attack the idea that essence can be understood in terms of real definition. As a matter of fact, this Aristotelian thesis has been the subject of a lack in popularity over the past century—especially due to Quine and the works of Kripke on *de re* necessity—but it has not, to my knowledge, been criticized *per se*. The idea that we can define things in the same way that we can define words is generally assumed to be uncontroversial.

Rather, the skeptic's criticisms that I will discuss target the idea that essence can fulfill the task of being the source of metaphysical necessity.

To begin with, Penelope Mackie expresses her concerns as follows:

I do not see how the notion of real definition [...] can do the job that is required of it in an essence-based theory of metaphysical modality. I do not see how it is possible to isolate a notion of real definition that will generate a conception of essence that delivers the result that essential properties are necessary properties [...], *unless* we appeal to modal notions—in particular, to the notion of metaphysical necessity—in explaining what the relevant notion of real definition is. It looks as if the account of essence in terms of real definition is intended to deliver a modal rabbit out of a non-modal hat. And I do not see *how* this can be done. (2020: 252, emphasis mine)

Mackie specifically targets Fine's own theory of necessity in terms of essence—*i.e.*, the reductive/identity theory. I think that her use of the term 'generate' as well as the 'how'-question she asks directly refer to a metaphysical gap, in the sense that she clearly finds that something is lacking in the essentialist's explanations, and I will elaborate this point a bit below.

In the previous section, I briefly mentioned that when facing a gap, the identity relation is not sufficient to support the kind of metaphysical explanation that we are after. Hence, something more (*i.e.*, another principle/relation/law/etc.) is needed in

this case to help bridge the gap between essence and necessity. This, in my opinion, corresponds exactly to the difficulties identified here by Mackie, because merely stating that essences are to be understood as real definitions and that necessary properties just are essential properties is not sufficient: why not schmessential properties? Or intrinsic properties? It seems that we do not have any sustainable ground to accept the identity of necessities with essentialities except for what the essentialist tells us: ‘that’s just how it is, period’. So something more needs to be said. I think the important point here is that Mackie is not saying that the view she criticizes is absurd or false. Rather, she is genuinely asking essentialists to give a substantial ‘story’ about their theory, because she simply finds it hard to understand or, simply put, ‘mysterious’ as to how it is supposed to achieve its goals.

Next, Carlos Romero (2019) raises similar worries and criticizes Bob Hale’s views on the matter. I will return to Hale in the next section, and for now Romero’s criticisms can be read as a general concern against Finean essentialism ‘broadly’ construed. Romero explains that essentialism “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” (2019: 5). He considers the following scenario:

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* is it 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*.

He continues:

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 is it that *a* is *necessarily F*.

(*Ibid.*: 6-7)

Let us appreciate that, in Romero's first quote, his remarks echo what has been said in the previous section. In particular, the gap implies that mere facts about the phenomenon in which we want to reduce another phenomenon to are not sufficient; in this case, mere essence-facts are not enough, according to him, to explain necessity-facts. Romero's main criticism is that the explanation does not get off the ground unless it is assumed that essences are necessary. Recall what has been said in the previous section regarding the 'is-ought'-gap: we can hardly infer an 'ought' from an 'is' unless we (tacitly) presuppose an additional normative premise or assumption together with/about the 'is'-claim. Similarly, what Romero points out here is that it seems implausible to have necessity 'popping out of nowhere' unless we (tacitly) presuppose that essence is itself necessary *and* that *this* modal status plays a substantial role in bringing forth the *explanandum*. If Romero is right, the idea that essence is the source of all necessity becomes hardly tenable because

essentialists just explain one necessity in terms of another: the only difference being that the *explanans* has an essentialist ‘flavor’.

Lastly, Jessica Leech raises another share of difficulties against Finean essentialism. She basically agrees with what Mackie and Romero say, and explains that it is not obvious that essence should imply necessity, nor that essence-truths are necessary either. She mentions that

once we reject an account of essence in terms of necessity and rethink our notion of essence anew, it becomes an *open question* whether we should expect essence to imply necessity in the way required by [NP]. The essentialist should not be swayed by a hangover from a view they reject. (2020: 5-6, emphasis mine)

Here, Leech complains that just because Finean essentialists reject the idea that essence can be defined in terms of *de re* necessity, it does not necessarily follow that a principle like NP is true. Of course, NP is true on the simple modal account of essence, but such an account (as well as sophisticated ones like the one I advocate) fall short of providing a definition (or reductive theory) of necessity since they rely on that very same notion to define what essence is. Moreover, Leech shares Mackie and Romero’s view about real definitions: “what something *is* does not tell us – absent further assumptions – what something *must be*” (2018: 320). So far then, real definitions state facts about things, but whether these facts are necessarily true is not something that is answered solely by grasping what the concept of a real definition is.

What these philosophers all have in common in their criticisms is that they do not see *how* and *why* essence—understood as real definition—should give rise to/be the source of metaphysical necessity. In sum, they all seem to agree on two points: first, that it is hard to understand how essences *simpliciter* could generate necessity; second, that *if* the modal status of essence is taken on board, *then* we can have a framework where necessity is passed from one set of facts (the essentialities) to another set of facts (the necessities), but that this jeopardizes the very idea that essence is the source of necessity. To compare with the gap in philosophy of mind, this would mean that if we presuppose that physical states also have mental features, then, of course, we can explain how particular mental states arise from physical states. However, this would not be a satisfying explanation at all of how the mental is generated out of the ‘purely’ physical since it is presupposed that the physical already has mental features. Similar reasoning applies to the normative and ‘is-ought’-gaps.

Thus, skeptics have strong reasons to think that there is a gap between essence and necessity. Their complaint is that essence does not, in and of itself, seem to be doing anything relevant that pertains to explanations of necessity unless it assumed that they are themselves necessary, and that this modal status plays a substantial role in bringing forth explanations of necessities. This is where the point of disagreement between essentialists and skeptics lies. What we need to do now is to answer the substantive question when facing an explanatory gap. In the present context, this amounts to the following inquiry: what is the characteristic trait of *necessities* that prevents a straightforward explanation of them in terms of *essence*? The obvious answer is that necessities are, by definition, *modal*. And according to

skeptics, this poses a problem for essentialists because they want to have (or reduce/ground) a modal phenomenon in terms of a *non*-modal one. What is important in this discussion is that it is the essentialists who owe the skeptic an explanation, and not the other way around. After all, essentialists are the ones claiming that essence is the source of necessity, and the skeptic merely asks how is that supposed to work. So the burden of proof is on the essentialists' shoulders.

Again, this situation is similar with the gap in philosophy of mind where the phenomenological aspects of mental states escape the explanatory scope of (neuro)physical states. In that sense, the essence-necessity gap is not 'cheap' like Schaffer's H<sub>2</sub>O gap. The reason is that, as soon as one attempts to describe necessity, one quickly realizes that it involves modal notions, and it seems that one can never get completely rid of them. This is a striking similarity in what all the authors above are saying: how is it even possible to discard all the modal terms and hold true to the view that essences *simpliciter* are the source of necessity? Or that for a characteristic *c* to be necessary to some *x* just is for *c* to be essential to *x*? I think that the comparison is even more vivid with the 'is-ought'-gap: we cannot infer that Mary ought not drink fifteen cans of Dr. Pepper a week from the mere fact that she actually does so, *unless* we presuppose that this is a *bad* (and rather unhealthy) habit. Similarly, the complaint of skeptical philosophers against essentialism consists in saying that they do not understand how a modal proposition can be derived from a non-modal one, unless substantial modal assumptions are made. Finally, we can compare the situation essentialists are in to the normative gap as well: one can see that it is essential to Socrates that he is human and yet genuinely wonder whether that is also necessary: 'I understand that

Socrates is essentially human, but I wonder whether he is also necessarily human'. These kinds of puzzlements strike me as genuinely legitimate and pointing towards a substantial difficulty, namely that mere *reduction* cannot meet the skeptic's demands for explanation.

It is now clear what kind of explanatory demand is required of any attempt at a reduction of necessity to essence. Since the philosophers I have quoted above do not explicitly mention whether they challenge essentialists on epistemic and/or metaphysical grounds, I will try to not misinterpret these authors. So what follows is my own interpretation of these problems. On the one hand, if the explanatory gap difficulty is a solely epistemic challenge, then the task for essentialists is to provide *reasons* to believe that essence gives rise to necessity—*i.e.*, *how* do we know that **NP** and all its multitude of instances are true, for instance. If, on the other hand, the explanatory gap is construed as a metaphysical difficulty, then the challenge for essentialists is to provide explanations regarding the *mechanisms* by which essence operates to bestow necessity on (a set of) propositions. Or, put differently, to address the skeptic's concern about whether essence and modality are related in the first place.

Either way, while it is legitimate for the skeptic—and anyone, for that matter—to ask why a principle like **NP** is true, I will solely focus on the metaphysical gap and seek an answer to the metaphysical question of *how* essence is supposed to give rise to necessity. The reason is that, in the case of essentialism, I do not think that it is possible to give a satisfying answer to the epistemic question without providing a detailed metaphysical explanation of the processes by which essence qualifies as

*the* source of necessity. I am not sure whether a similar reasoning applies to the normative/phenomenological cases, and I will not investigate these questions here, one of the reasons being that these theories presumably try to explain a subjective feature in terms of an (empirically) objective one. In fact, it is an open question whether there is such a thing as objective goodness and, quite obviously, phenomenological states are subjective in nature and depend on there being sentient beings to experience them. This means that these cases of explanatory gaps concern, in great part, how our human lives are/can be governed (assuming that there is no objective goodness). Consequently, there is some threat that the debate is first and foremost epistemological and about our cognitive abilities, since the questions at stake are, by definition, *about* what being human is; we could imagine worlds without normative principles, and without phenomenological states. The essence/necessity gap, on the other hand, is substantially different since these notions include, but do not *essentially* involve, human beings, their actions, and their phenomenological states. What is required from the essentialist is to bridge a gap between *two objective phenomena* that are present regardless of whether sentient beings exist to think about or experience them; it is an investigation about the ultimate nature of reality and how it is structured, and it is not contingent on there being humans with emotions and/or notions of what they perceive as being (objectively) good or evil, for instance.

In saying this, my point is that essence and metaphysical necessity are supposed to be absolute and objective matters. If essence *is* the source of necessity, as essentialists claim, then it is what objectively governs and shapes not only *this* world, but *all* worlds—the myriad of all possibilities that are ‘out there’. So unless

we have an account of *how*—metaphysically speaking—essence achieves such a task, it will be hard to have reasons and to be justified in believing that essence generates necessity. Similarly, objective empirical evidence gives us reason to believe that water is constituted of H<sub>2</sub>O molecules. And this is quite different from the phenomenological gap: since we cannot have—at least for now—direct and objective access to a sentient being’s mental states, it is hard to start with the metaphysical question and, say, provide empirical evidence (*i.e.*, scientific and metaphysical explanations) of the underlying mechanisms that govern such interactions. Rather, *even* if the philosophy of mind gap is a substantial difficulty, it is first and foremost an epistemic issue precisely because of the limited access we have of the mental. If mental states are identified with (neuro)physical states, the question remains open as to why pain/pleasure feels the way it does.

Also, identifying mental states with (neuro)physical states leaves psychophysical interactions completely mysterious. In comparison, the gap between essence and necessity is slightly different: if we identify/reduce necessities to essentialities, we thereby provide no explanation whatsoever of the mechanisms that govern these notions. In fact, this amounts to saying that there is *no such interaction*. And as opposed to the phenomenological gap, such forms of identification/reduction do not leave something mysterious like psychophysical interaction out of the picture, since necessity is not supposed to interact, in turn, with essence. That is, if an essence-fact, *E*, explains  $\Box p$ , it does not happen in turn that  $\Box p$  plays an explanatory role towards *E* or some other essence-fact *E'*. However, if mental state *M* is reduced to physical state *P*, then there might be a situation where the phenomenological aspects of *P* cause another physical event *P\**. All of this is to

say that reducing necessity to essence leaves us without any possibility to account for the mechanisms by which essence operates: it is to say that for  $p$  to be necessary just is for  $p$  to be true in virtue of the essence of some  $x$ . The reasons for believing that this is true have to be taken for granted: to be necessary is to be true in virtue of the essence of some  $x$ , period. This is explanation by stipulation, and therefore no explanation at all.

It seems clear to me at this point that the best and only way essentialists can answer the metaphysical question is by opting for a *non-reductive* account of necessity. That is, to say that necessity is *grounded* in essence, as on Hale and J. E. Lowe's views. This leaves room for discussion and, most importantly, *explanation*. We can now imagine that there are two different and *distinct* realms: the realm of essences and the realm of metaphysical modality. Had essences failed to be there in the first place, the modal realm would not be. By analogy, we can imagine how, say, the 'moon-realm' contributes/influences tides in the 'ocean/shores-realm': both realms are different, but connected, and an explanation can be given as to how and why the moon influences high tides and low tides. I hope the example is not too far fetched, but I think it helps to see what the essentialist should be after: giving an explanation of *how* the realm of essences gives rise to the realm of metaphysical modality, *how* both are connected. This can be done via proof, argument, logic, together with some kind of thought experiment. Last, I think it is important to observe that what essence is—*i.e.*, how it is construed—might also be an obstacle that prevents a straightforward solution to the gap problem. I now turn to the question of how and whether the gap between essence and necessity can be bridged by essentialists.



## 2.III On bridging the gap between essence and modality

Essentialists tell us that the essence of a thing can be understood as the real definition of that thing. The purpose of a real definition is to state what a given thing *is*, but the problem for essentialists is to explain *how* the connection is made with necessity and/or to explain what—in this case—is the special feature of real definitions that enables them to bring forth necessity.<sup>28</sup> As explained in the first chapter, there are two different views about essence understood as real definitions: one is the ‘pure’ Finean reductive approach where necessity is defined in terms of essence, and the other is Hale’s non-reductivist theory according to which necessity is *grounded* in essence. I will try to outline what the possible answers are for each of these views, starting with Finean reductionism.

A bad answer from the essentialist would basically consist in saying that real definitions are themselves to be defined as sets of propositions that state the necessary properties of the things they are about.<sup>29</sup> However, this cannot count as an explanation at all, since it is nothing but an *ad hoc* stipulation. To illustrate, it is as if someone was asking why some birds migrate, and was told in response that

---

<sup>28</sup> There is also a related difficulty in the sense that we have no clear ‘recipe’ regarding which features are essential to some given *x*. For instance, Fine’s followers seem to agree with the intuitive idea that Socrates is essentially human, but this is not something Mackie (2006) agrees with.

<sup>29</sup> We can be doubtful about the idea that notions of real definition should themselves be defined. It is one thing to say that a definition states the necessary and sufficient conditions for a thing, *x*, to be an *F* (e.g., for *x* to be a mammal, or for *x* to be Socrates), but quite another to say that definitions are themselves defined in terms of what states the necessary properties of *x*. We can imagine notions of real definitions that list contingent properties as well. And a full definition of Socrates will not only include necessary properties like being human, but also (arguably) contingent properties like being Greek, having *x* amount of hair on his head at time *t*, and so on.

this is so because (part of) what it is to be such birds is to migrate; or, to stick with a more familiar example, it is the same situation as when the essentialist says that necessities just are essentialities because it is so. There remains a deep dissatisfaction with such answers; it is, in fact, as if the interlocutor was trying to dodge the question rather than answer it. In general, then, I am inclined to say that the kind of ‘whether’, ‘why’, and ‘how’-questions we are after—*i.e.*, whether/why/how does essence explain necessity—can never be properly answered with a definition.

If I am right about this point, then the gap is an immediate problem for reductionists like Fine. Remember that Fine spells out his views on essence and metaphysical necessity as follows:

The metaphysically necessary truths can then be *identified* with the propositions which are true in virtue of the nature of all objects whatever. (1994: 9, emphasis mine)

This is not merely a reductive thesis, but clearly a definitional one, as Fine clarifies in the following passage when comparing his view to Husserl’s:

I do not follow [Husserl] in treating the necessity in question as a priori or in taking the essences of things to be universal; and he does not follow me in treating the account as a *definition* of one notion in terms of another. (*Ibid.*: footnote 1, emphasis mine)

As readers of Fine, we cannot help but be dissatisfied, since in the aforementioned paper as well as in some of his other works, he does not provide us with any reason to think that essence should be the source of necessity, nor does he explain how essence happens to be somehow tied up with necessity in the way he thinks it is. It therefore strikes me that the relation between the two notions is unjustified and must be taken for granted. In his own words, again, Fine says that

there is a connection between the two concepts. For any essentialist attribution will give rise to a necessary truth; if certain objects are essentially related then it is necessarily true that the objects are so related. (*Ibid.*: 9)

There is not the slightest trace of an explanation nor of an argument in these passages as well as in Fine's other papers. In response, I would like to say something that echoes Leech's comments in her 2020 article, namely that as soon as the modalist view of essence is rejected, the idea that essence and necessity are somehow related should entirely be thought anew, argued for, as well as defended on solid grounds, but not anchored into mere stipulations. In fact, no one accepts the physicalist theory of the mind at face value (except for some physicalists): if someone said 'all mental events just are physical events', the natural reaction would be something like 'really? how come? can you elaborate?'; for over 50 years now, physicalists brought all kinds of arguments to sustain their view, and this is something that is entirely absent from the neo-essentialist's literature.

In brief, what Fine says in the above quote looks like a result, a conclusion, but we ought to wonder how these are reached. Finean essentialism first originated from arguments against the idea that essence can be reduced to necessity, since our intuitive judgments about certain cases reveal that the view is not sufficiently fine-grained. Then, a quick step is made towards the idea that the relation should be thought of the other way around. But why? The fact that Fine grounds his view on the Aristotelian model of real definitions is not, in and of itself, a sufficient reason to endorse the view that essence is the source of necessity. For just because we once thought that essence could be defined in terms of *de re* necessity does not mean that, as soon this view is rejected, there is (or even *ever* was) a link from essence to necessity.

But let us (re)consider modalism for a moment: is there not something intuitively and at least partly right in saying that if Socrates is necessarily human, then he is essentially human? In that particular scenario, whatever the necessary properties of *x* are, they are also essential to *x*, because having a property essentially consists, in part, in having that property unconditionally—*i.e.*, in all possible situations: it is a property that makes a thing the very thing it is, and if such a property were lacking, that thing would be different. That is, in my opinion, the general intuitive idea behind modalism as intended in the first place, and why it does not seem that a gap between necessity and essence arises. What is crucial to note here is that we stay *within* the modal realm with these explanations: the essential properties coincide with the necessary properties; they are one and the same. (Whether this is fine-grained enough to theorize about the essential properties of some *x* is not the concern here—*e.g.*, Fine's counterexamples involving {Socrates}.) Rather, the

point is that, generally speaking, when we are in the business of talking about essence, we somehow want to secure the idea that entities cannot *not* possess their essential properties, which I take to be the core intuition behind any understanding of essence, regardless of any independent consideration. And only *then* can we start theorizing about whether the necessary properties we have singled out are accurate answers to ‘what’-questions (*i.e.*, what is Socrates?).<sup>30</sup>

To illustrate, imagine yourself explaining to a non-philosopher what an essential property is. At some point, I bet that you will have to say something closely related to permanence, invariability, unchangeability, etc.—which are all inherently modal notions that echo what we, philosophers, like to call ‘metaphysical necessity’ in our technical jargon. If this is not how essentialism is thought or conceived of at least partially, then I truly do not know what it is nor how it can be a candidate for being the source of necessity. To that effect, then, I fully sympathize with the skeptical philosophers discussed in the previous section.

To summarize, it does not seem that we need a prior understanding of essence to grasp what necessity is. Rather, it appears to be quite the opposite: it seems impossible to understand what essence is without, at some point, involving modal vocabulary—regardless of whether modalism is true and similar kinds of reduction are possible. So when Fine says that since Socrates is essentially human, then he is necessarily human, it immediately prompts the gap question. The reason, I believe, is quite simple. In this particular scenario, the modal sphere is—as opposed to on

---

<sup>30</sup> This is actually the core idea behind sophisticated modal accounts such as those due to Bovey (2021), Correia (2007), Denby (2013), and Wildman (2013). According to these views, the modal part is a non-negligible element of what essence is.

the classical modalist view—defined into something that is non-modal. (Let me stress here again that one cannot define a thing in terms of the very thing that must be defined, which is absurd.) The reductive essentialist’s problem is therefore twofold.

*First*, as opposed to modalism, pure reductionists cannot make sense of the very idea that essence is *itself* a modal phenomenon. The reductive/Finean approach forces its defenders to embrace the claim that essence is not modal.<sup>31</sup> This is a thorny matter, for the reasons mentioned in the paragraph before last but not only. For if a truth, *p*, is necessary and has its source in an essentialist truth *E*, does it not also mean that *p* has and *should* have *E* as its source across all possible scenarios, and that *E* is *eo ipso* metaphysically necessary as well? In the passage quoted above, Fine himself says that *surely* there is a relation between essence and necessity, but my point here is that *surely* this relation should hold no matter what: whenever I see *p*, I should see *E*, and *vice versa*. So if metaphysical necessity is to be defined in terms of another phenomenon, the said phenomenon should at the very least make sense of the idea that metaphysical necessity is absolute and holds across both actual and non-actual situations; and I am not sure how this can be done without assuming that essence is itself necessary—that it is itself established and available as a source across all possible scenarios.

---

<sup>31</sup> Essence is supposed to explain the strongest kind of necessity. If essentialists appeal to another ‘stronger’ kind of necessity that pertains to essence only, this strikes me as *ad hoc* and unjustified. What could be more absolute than metaphysical necessity? Once we grasp what this concept is, it is hard to make sense of something even stronger.

Thus, under Fine's reductionist view it is as if essentialist claims were true in all situations, but that this had nothing to do with modality at all. Rather, it would be thanks to some kind of 'coincidence', and the very fruit of that coincidence—that is, its consequence—happens to be metaphysical modality. This looks like a rather fragile framework. But most importantly, it strikes me that a reductive construal of essentialism is at odds with a plausible metaphysical thesis: *modal universalism*. According to this view, every truth is either contingent or necessary. It is, in a sense, similar to the semantic principle of bivalence in classical logic. So regardless of whether necessity can successfully be reduced to essence or not, the question of what the modal status of essentialist truths is will remain open. And it seems that reductionists cannot answer such a question on pain of a regress: if essences are necessary—as I think they should be—then they themselves reduce to some essences, which are in turn necessary and so on *ad infinitum*.<sup>32</sup> Under such circumstances, I believe that the prospects of a successful reduction are dim.

So for Fineans, essence has to come first 'at all cost', and independently of modal notions. I must say that this is rather strange, because the modal realm comes out of something completely 'foreign', and my intuition here is that, roughly put, 'dogs do not make cats'. In other words, I side here with the skeptic and wonder *how* can essentialist truths generate something as strong as metaphysical necessity without being themselves governed by some kind of 'law' that is at least as strong as metaphysical necessity (if not stronger)? It seems that essence can never be the sole source of modality without being itself entangled with modality. Necessity resists an analysis in purely essentialist terms, and in rejecting the idea that essence

---

<sup>32</sup> More on this point in chapter 3.

reduces to necessity, Fine throws the baby out with the bathwater by defending a view where essences have to be non-modal. For unlike modalism, truths that hold essentially are no longer (and cannot be) a subset of the necessary truths, and it is therefore not only counterintuitive to think that essences are not necessary, but also hard to make sense of the idea that a non-modal phenomenon can be the source of necessity.

This takes us precisely to the *second* problem, which is nothing but *the* gap. For the situation essentialists are facing is similar to the mind-body problem: reduction seems doomed to fail, and just as it is hard to wrap our heads around the idea that pure physical events can give rise to mental and subjective events, it is hard to make sense of the idea that a non-modal conception of essence can give rise to something as special as metaphysical necessity.

At that point, I am convinced that the reductive view of necessity to essence is doomed to fail and that it can never meet the explanatory demands of the skeptic. One might be tempted to say that it is a stalemate and that one merely has to pick a side: either you think that essence generates necessity without further questioning, as if it were ‘self-evident’, or you just cannot accept that thesis without additional premises, explanations, and reasons to believe it. However, I do not think that such matters are as easy as I have just sketched. In fact, the essentialist cannot be put on par with the skeptic in this situation as if it were an even bet. For the essentialist is the one who *first* claims that essence is the source—and therefore explains—necessity, while the skeptic is simply asking for justification, reasons, and—most importantly—detailed metaphysical explanations of how essence achieves such a



task. I take the skeptics' questioning to be perfectly legitimate, as the burden of proof does not lie on their shoulders, but on the essentialists'. Since the reductionist does not have the resources to bridge the gap, as I have just tried to argue, we have to look elsewhere in the literature to find answers.

The other candidate framework to consider is Hale's, who, just like Fine, thinks that essence is the source of necessity, but who argues that no reductive account of necessity can ever be achieved: be it in terms of worlds, essence, or whatever other kind of truths or entities there might be. To that effect, Hale is a non-reductive essentialist who argues that necessity is *grounded* in essence. Ontologically speaking, this means that reality is both composed of, and structured by, (at least) two different layers: at a (relatively) fundamental level lie essence facts from which necessity facts are derived.

There are several reasons why Hale is not a reductionist and I will focus on the one I consider to be the most relevant. As opposed to Fine, Hale is first and foremost interested in addressing Blackburn's dilemma (1986). Briefly, this dilemma is a challenge to any theory who aims to identify the source of necessity: either the source is contingent, in which case had the source failed to obtain, the necessities would no longer be necessary; or the source is itself necessary, in which case we face the threat of a regress because if, say,  $\Box p$  is explained by  $\Box q$ , we now have to explain (*i.e.*, find the source of)  $q$ 's necessity.<sup>33, 34</sup> Hale thinks that one necessary

---

<sup>33</sup> Another problem would be circularity:  $\Box p$  is explained by  $\Box q$  which, in turn is explained by  $\Box r$ , and eventually we would arrive at point where the *explanans* of  $\Box r$  is  $\Box p$ .

<sup>34</sup> For the purposes of the present investigation, I will only discuss parts of the dilemma when relevant. For a detailed discussion, see Cameron (2010).

condition for a phenomenon to be the source of necessity is for the said source to be itself necessary. Hence, Hale not only sits on the necessity horn of Blackburn's dilemma, but he also finds himself in a position where he can answer the modal universalist's question, namely: what is the modal status of essentialities. Next, Hale's first step to address the necessity horn consists in defending the claim that even though essence *is* the source of necessity and is itself necessary, the modal status of essence *cannot* be explained. In other words, essentialities are fundamental metaphysical necessities for Hale. This means that, for him, the set of necessities is divided between the fundamental necessities and the derivative necessities: essence is the source of *all* and *only* the derivative necessities, nothing else. This is a much weaker thesis than what essentialists like Fine initially wished to defend, and also weaker than other frameworks like Lewis's, who attempts to convey explanations for all necessities, without exception.

Clearly, it would be best for essentialists if they could simply find a way to maintain that essence is necessary (to avoid the difficulties that reductionists face), and that essence is the source of *all* necessity—granted that the gap problem can be solved. In the next chapter, I will explain why I think essentialists cannot defend this thesis universally, and why, no matter what, they have to opt for a relativized/restricted framework like Hale's. But before engaging this topic, let us see what Hale is trying to accomplish with his non-reductive view, and whether he has a way to bridge the gap between essence and necessity.

So far, we might think that Hale's explanations of necessity look like this:

$$[\Box E(x)p] < [\Box p]$$

(Read: The fact that it is necessarily essential to  $x$  that  $p$  grounds the fact that  $p$  is necessary.)

We are not really making any progress if this is the case, because it does not look like the essence *simpliciter* is doing any particular work in explaining  $p$ 's necessity. (What Hale's strategy does, however, is to stop the demand for explanations of the necessity of essence facts since they are fundamental.) Hale would agree with this point, since he thinks that the above explanatory schema corresponds to a *transmissive explanation* of necessity, one where the necessity of the *explanans* is 'transferred' to the *explanandum*. Hale himself explains that

[t]he transmission model offers no prospect of an explanation of necessity in general. An explanation of necessity in general need neither be, nor provide for, an explanation of each and every particular necessity. What is required, rather, is to explain why there is any necessity at all—we may, and I shall, take this to amount to the requirement to explain why, assuming it to be so, there is at least one necessary truth [...]. (Hale 2013: 96)

What is important to keep in mind is that Hale is interested in answering one specific question, namely 'why is there *any* necessity?' as opposed to 'why is *this particular* truth necessary?'. This subtle point is intimately tied to Dummett (1959) and Blackburn's considerations about the source of necessity. Dummett says that

one of the main philosophical problems about necessity is to find out what its source is and, following in his footsteps, Blackburn explains that philosophers searching for the source of necessity are not after a *local* proof of why a particular truth is necessary, but rather are after what he calls ‘the whole area’ (1986: 53). This ‘whole area’ is what Hale is referring to in the above quote when he talks about ‘necessity in general’.<sup>35</sup> So what is important here is that, as opposed to Fine, Hale approaches the question of the source of necessity entirely through the lens of Blackburn’s dilemma. Hence, if Hale opts for the transmission model of necessity sketched above, it is impossible for him to succeed, because in order to explain why there is a necessity, he should not presuppose what he is meant to explain, namely: necessity. To clarify, on the transmission model of explanations of necessity, there *is* what Blackburn calls a ‘bad residual must’, because we have not localized what the source of necessity is: we merely say that some truth owes its necessity to another necessity. To better understand what is the problem at stake, let me mention one useful analogy from Cameron. He explains that, on the transmission model of explanations of necessity, we have

relied on the fact that there are necessary truths to explain why there are necessary truths; it is as if I have attempted to explain why there is something rather than nothing by saying ‘well my parents gave birth to me, which is why I exist; hence I exist, hence there is something rather than nothing. (2010: 6)

---

<sup>35</sup> See also Cameron 2010.

In that example, we only know why a *particular* thing exists—*i.e.*, we have a ‘local’ explanation of why Cameron exists, but we do not know why there is something rather than nothing—which amounts to a *general* question about the very nature existence, its source.

So, according to Hale, what is required for the non-reductionist essentialist to succeed is that the *explanans* of the necessity of some truth makes *no mention* of necessity whatsoever. For it is clear that if one explains  $\Box p$  in terms of  $\Box E(x)p$ , one is left with the latter necessity being unexplained and the source of necessity unidentified. But if, as Hale claims, essence is both necessary and fundamental, and if transmissive kinds of explanation are doomed to fail, then how can he avoid Blackburn’s ‘bad residual must’? Hale’s answer is that essence provides *non-transmissive* kinds of explanation. In this particular scenario, this means that even though essences are necessary, it is their *truth*—and their truth *alone*—that bestows necessity on their prejacent proposition. Schematically:

$[\Box E(x)p], [E(x)p], \text{ and } [E(x)p < \Box p]$

So what this means is that only  $[E(x)p]$  is relevant to  $p$ ’s necessity. With this kind of explanation, Hale does not presuppose the necessity of essence in order to explain why  $p$  is necessary, and he supposedly gives an answer to the question of why there is at least one necessary truth. However, it is obvious that this view faces substantial problems, the prominent one being that if the modal status of essence is —and *should* be—explanatorily irrelevant, we have no answer regarding how essence generates/grounds necessity. We are told that essence grounds necessity,

but it remains a mystery why and how it does so without having a modal component involved in the process: how can something bestow a feature that it does not possess? Hale made a step forward by providing an argument to the effect that a necessary condition for something to be the source of necessity is for it to be itself necessary. However, even if Hale finds a way to address the ‘bad residual must’ of the necessity horn of Blackburn’s dilemma (and this is actually debatable), his framework ultimately does nothing more than Fine’s reductive view as far as the gap is concerned. In fact, on Hale’s view we arrive at exactly the same situation as the Fineans are in if they take modal universalism, the necessity of essence, and Blackburn’s dilemma seriously.

So far, then, the two main essentialist frameworks we have just discussed stand on the exact same ground when facing the gap. But before moving on to the next chapter, I would like to highlight that, just like Fine, Hale gives us no reason to think that essence—as opposed to something else—is the source of necessity: this claim is taken at face value and there is no substantial argument provided to sustain it. Thus, questions like ‘why essence rather than something else?’ remain unanswered. And again, what is even more problematic is that when we ask how does essence bestow necessity to its prejacent proposition, it is hard to imagine how the necessity of essence is not somehow involved. I will return to this point later in chapter 5, where I will discuss a sophisticated version of Hale’s non-reductive essentialism.

To conclude this chapter, essentialists cannot be reductionists because this view cannot make sense of the idea that, intuitively, essences are necessary. Yet, if the

kind of reduction essentialists are after is supposed to succeed, essences have to be non-modal. The second point against reductionists is that even though they say that necessity reduces to essence, it remains an open question whether essences are necessary. The viability of the reductive framework gets even weaker in light of the gap problem and, at this point, it is clear that the best essentialists can hope for is a non-reductive framework like Hale's. However, as we have just seen, this view cannot possibly bridge the gap between essence and necessity, since its success (in overcoming the necessity horn of Blackburn's dilemma) rests entirely on the premise that the *truth*—and *not* the necessity—of essence is what grounds necessity. To that end, the reductionist and the non-reductionist are on the same page and differ only in their commitment to addressing other difficulties and on how they frame the theory. As explained in the first section of this chapter when discussing the 'is-ought' gap, from the mere fact that it is essential to  $x$  that  $p$ , it does not follow that  $p$  is necessary—unless we make (tacit or explicit) modal presuppositions regarding the essentialist claim.

Perhaps there is a way to address these problems if essence is understood in a different way—*i.e.*, not in terms of real definition. However, I will turn to this question in chapter 4, where I will consider the idea that essence is intimately tied to identity. First, I want to focus on another problem that I have merely touched on in this chapter, namely: the necessity of essence. While the explanatory gap challenge poses a threat to the intelligibility of the sourcehood thesis, the problem I will now tackle focuses on the extensional adequacy of the essentialist theory of necessity. If essentialists fail to have the resources to address the question of explaining the necessity of essence—as I shall argue—this will give us even more

reason to turn towards another construal of essence-truths. So in the next chapter, I will focus on non-reductive essentialism and whether this view can, as opposed to what Hale argues, explain *all* necessities. One of the downsides of Hale's view is that it is a weaker thesis than Fine's in the sense that it admits of fundamental necessities that cannot be accounted for. If it is possible to maintain the idea that essence is the ground of all necessity, it might provide us new grounds for thinking of whether there is a way to bridge the gap.



## Chapter 3: On the necessity of essence

In this chapter, I want to bring a second substantial difficulty against essentialism.<sup>36</sup> My goal is to figure out whether Hale is right in thinking that the necessity of essences must be fundamental on pain of being confronted to vicious circularities and/or regresses.<sup>37</sup> If I am right, my arguments below will reinforce the idea that not all necessities can be explained in terms of essence, and that essentialism must undergo substantial revisions in order to uphold the idea that it can explain all necessities.

The plan is as follows. In section 3.I, I set the stage for the main discussion with a characterization of the sourcehood thesis and a principle according to which essences are necessarily true. Next, I introduce an explanatory structure that will be the main point of discussion in the remaining sections. In section 3.II, I give a first assessment of the explanatory structure, and I focus on the nature of the essence-facts involved in explanations of necessities to show how they are explanatorily connected with one another. Next, I defend a principle that bridges explanations of necessity with explanations of truth, and I put forward the main circularity objection against explanations of the necessity of essences in terms of essence. In

---

<sup>36</sup> This chapter has been published in *Philosophical Studies* (2021) and the format has been adapted for the present discussion. Slight modifications have been made but the structure and the arguments remain faithful to the published version.

<sup>37</sup> See Hale (2013: 158). In fact, Hale does not *prove* that there are such problems at stake for essentialism.

sections 3.III and 3.IV, I outline a series of solutions to the difficulties raised in the main discussion, and I conclude the chapter by briefly presenting my own take on these issues.

### 3.I The necessity of essence

As a first step, I want to characterize **Source** more precisely than in the previous chapters for the sake of readability and clarity:

**Source:** if  $\Box p$ , then  $\Box p$  because  $\exists x(E(x)p)$ .<sup>38</sup>

(Read: if it is metaphysically necessary that  $p$ , then this is so because there is some  $x$  such that it is true in virtue of the nature of  $x$  that  $p$  (or, equivalently, such that  $x$  is essentially such that  $p$ )).

Two important remarks are in order. First, the propositional variable ‘ $p$ ’ is construed as being universally quantified, so as to capture the idea that essence is the source of all metaphysical necessities. Second, I read the ‘because’ locution as denoting *metaphysical explanation*. What I mean by that term is a non-causal form of explanation that obeys priority and dependence constraints.<sup>39</sup> The reason is that I

---

<sup>38</sup> Let me stress again that I will solely focus on explanations of necessity that appeal to the essence of a single entity. Again, some necessities are true in virtue of the nature of a plurality of entities—*e.g.*, cases of numerical distinctness. For simplicity, I will not give another name to the sourcehood thesis, this will not affect the following discussion.

<sup>39</sup> See Maurin (2019) for the general notion of metaphysical explanation, as well as the dependency constraint.

want to remain as neutral as possible to properly conduct the following investigation—*i.e.*, without any particular commitment to, say, grounding or any other specific kind of explanation. With respect to the essentialist framework, this means that essence is metaphysically *prior* to necessity, and that the latter depends on the former. I also assume that metaphysical explanation possesses the following properties: *asymmetry*, *irreflexivity*, and *transitivity*. Importantly, my discussion is primarily tailored around the assumption that ‘metaphysical explanation’ is an ‘umbrella’-term under which all specific kinds of metaphysical explanation—*e.g.*, grounding—can be unified. I will flag in due time whether some of my arguments rely on a particular kind of metaphysical explanation and, in section 3.III, I will discuss the option of distinguishing between different kinds of explanation in order to explain how some difficulties might be overcome.

Next, among the philosophers who have followed in Fine’s footsteps, there is a conjecture according to which essence-truths are metaphysically necessary (*cf.* Hale, 2013, 2018; Lowe, 2012; Van Cleve, 2018; Wallner, 2019).<sup>40</sup> For it would be odd to say that essence-truths could have failed to obtain or that the nature of, say, Socrates could have been different. Moreover, as we have seen in the previous chapter, essentialists might have others reasons—such as Hale’s—to embrace the idea that essences are necessary. I contend that the necessity of essence-truths is, so far, fairly uncontroversial, and I will not challenge this claim in that chapter. In

---

<sup>40</sup> It is indeed a conjecture since there is not—to my knowledge—a successful argument to the effect that essences are necessary. One attempt is made by Hale (2013), but it presumably fails as Romero (2019) and Leech (2020) argue. See Wallner and Vaidya (2020) for a reply. However, it is possible to prove the necessity of essence—as Leech (2020) demonstrates—if one opts for Correia & Skiles’s (2019) ‘generalized identity’ essentialism, and I will discuss these views in chapter 4.

order to systematize my explanations, I capture the idea that every essence-truth is necessary with the following principle ('Necessity of Essence'):

$$\mathbf{NE}: E(x)p \supset \Box E(x)p.$$

Since **Source** holds for all metaphysical necessities and since, given **NE**, essences are metaphysically necessary, there must be an explanation of the necessity of essences in terms of essence. More precisely, if  $\Box E(x)p$ , then there is some  $y$ , such that  $E(y)E(x)p$ .

There are two possibilities: either  $y = x$  or  $y \neq x$ . The former option appeals to an iterated essence with an identity between the bearers:  $E(x)E(x)p$ . For instance, the essence-truth 'Socrates is essentially human' belongs to Socrates's very own essence. The latter option, however, is trickier, since it amounts to an appeal to an iterated essence without identity between the bearers:  $E(y)E(x)p$ . This requires an investigation on its own to find out which (if any) entity is such that it is essential to it that it is essential that Socrates is human (to continue with the same example). I will start my investigation by focusing on iterated essences with bearer identity to illustrate my points. Discussing this option first will allow us to have a better understanding of what is at stake in explaining the necessity of essences, as well as identifying which challenges essentialists face. I will turn to iterated essences without bearer identity in section 3.IV.<sup>41</sup>

---

<sup>41</sup> Glazier (2017) provides arguments to the effect that we can never have essence-truths of the type ' $E(y)E(x)p$ ' (where  $y = x$  or  $y \neq x$ ). However, I will ignore this for the sake of the argument and return to it later on. So far, the point of the discussion is to introduce how essentialists can provide systematic explanations of the necessity of essences.

With the assumption that  $y = x$ , **Source** and **NE** are jointly sufficient to generate the following structure of explanations:<sup>42</sup>

**Structure 1 (S1)**

4.  $[E(x) \dots E(x)p] < [\Box E(x) \dots E(x)p]$ ;
3.  $[E(x)E(x)E(x)p] < [\Box E(x)E(x)p]$ ;
2.  $[E(x)E(x)p] < [\Box E(x)p]$ ;
1.  $[E(x)p] < [\Box p]$ .

(Here, ‘<’ denotes metaphysical explanation.) The generation of **S1** starts with the explanation of one necessity in terms of an essence truth (line 1), and given **NE**, that essentialist truth is in turn necessary; per **Source**, that necessity is explained in terms of an iterated essence truth (line 2), which is in turn necessary, and this process repeats itself *ad infinitum*.

The way **S1** is structured prompts the question of whether it is—in some way or another—viciously circular and/or regressive. Reasons to think that this is the case greatly depend on whether there are—in addition to explanations of necessity in terms of essence—explanations among necessities and/or among essences, respectively. For instance, **S1** would be viciously circular if some facts eventually explained themselves, and viciously regressive if the success of some explanations

---

<sup>42</sup> Variants of this structure can be found explicitly in the works of Hale (2018), Van Cleve (2018), and Wallner (2019).

depended on infinitely many steps.<sup>43</sup> My goal in the following section is to provide detailed discussions of these issues, as well as important points of clarification. Doing so will help determine whether **S1** involves substantial difficulties of some kind.

### 3.II Assessment of S1

With what has been said in the previous chapter, one might immediately notice that it is possible for the necessity of essences to play an explanatory role in **S1**. However, for simplicity I will simply assume in the following discussion that essences can explain necessity, that there is no explanatory gap at stake, and that the necessity of essence is explanatorily inert. I will return to the role of the modal status of essence in explanations of necessity much later in chapter 5, when it will be time to draw conclusions from our discussions.

Let us now turn to the question of whether explanatory connections hold among essences in **S1**. Here, I contend that two options must be rejected: (**α**) no essence is explained by any other essence in **S1**, and (**β**) every essence in **S1** is explained by the essence directly above it—*e.g.*, ‘ $E(x)E(x)p$ ’ is explained by ‘ $E(x)E(x)E(x)p$ ’. In order to explain why neither of these options is sustainable, a proper investigation

---

<sup>43</sup> Let me stress that we should not think of essentialist truths in **S1** as entities—*e.g.*, as the essence of  $x$ , the essence of the essence of  $x$ , etc. For, as Lowe explains, this construal leads to an infinite regress that, “at worst, would be vicious and, at best, would appear to make all knowledge of essence impossible for finite minds like ours” (2008: 38–39).

regarding the ‘nature’ of the essences in **S1** is required, as well as important considerations about the direction of explanations, to which I now turn.

The first important point is Fine’s distinction between the notions of *constitutive* and *consequential* essence (1995):<sup>44</sup> the constitutive essence of an entity,  $x$ , contains the propositions that are directly definitive of  $x$ , while its consequential essence contains the propositions that are obtained when the constitutive notion is closed under logical consequence. For instance, it is constitutively essential to Socrates that he is human, but consequentially essential to him that he is human or a fish. Focusing on **S1**, ‘ $E(x)p$ ’ is a claim of constitutive essence since ‘ $p$ ’ is directly definitive of  $x$ —*i.e.*, about  $x$ . However, iterated essences with bearer identity such as ‘ $E(x)E(x)p$ ’ are not of the constitutive kind. For one thing, this claim cannot be directly definitive of  $x$  since it speaks about the essence of  $x$  and not about  $x$ . For another, Fine developed a Logic of Essence that governs the consequential notion of essence. In this work, Fine proves the following theorem:  $E(x)p \supset E(x)E(x)p$  (1995: 255). Thus, ‘ $E(x)E(x)p$ ’ is a claim that *derives* from ‘ $E(x)p$ ’ and that belongs to the consequential essence of  $x$  (and the same is true of all the other iterated essences in **S1**).

Given this important distinction, it is pretty clear that not all essences in **S1** enjoy the same metaphysical (and probably also ontological) status—for instance, as if each of them were entirely unrelated to the others. To pump intuitions, structures like **S1** are not, as it were, ‘discovered’. Rather, they are *generated* (as explained in

---

<sup>44</sup> Fine offers several ways to understand the distinction, sometimes taking one of the two notions as definable in terms of the other or *vice versa*. Here, however, I take no stance on such debates and solely rely on an intuitive way of spelling out the distinction.

the previous section). The crucial and decisive step in this process, then, is when appeal to ‘ $E(x)E(x)p$ ’ is made to explain ‘ $\Box E(x)p$ ’. To that effect, I contend that this iterated essence finds its source in—and is thereby metaphysically explained by—‘ $E(x)p$ ’.<sup>45</sup> Importantly, I am not saying that because ‘ $E(x)p$ ’ is the logical source of ‘ $E(x)E(x)p$ ’, it is therefore the case that ‘ $E(x)p$ ’ is also the metaphysical source of ‘ $E(x)E(x)p$ ’. Both claims are, in my opinion, true, but they are true *independently* of one another and they overlap in the present case. This is similar to the situation where ‘ $A$ ’ and ‘ $B$ ’ are not only the logical source of the conjunction ‘ $A$  and  $B$ ’, but also the metaphysical ground of the fact [ $A$  and  $B$ ]. The point, then, is that even if there is an entailment at stake, as Fine’s theorem establishes, it is also the case that, metaphysically speaking, iterated essences are generated from a constitutive (and non-iterated) essence: ‘ $E(x)p$ ’ is prior to ‘ $E(x)E(x)p$ ’, and the latter depends on the former. In fact, it would be a rather strange metaphysical framework if infinitely many essences like the ones in **S1** were independently ‘wandering’ out there, and if, despite their structural similarity, there were no explanatory way to connect them with one another.<sup>46</sup> We can clarify the explanatory connection at stake with the terminology of Correia (2012): consequential essences are derivative, because they admit of further explanation in essentialist terms (*i.e.*, it is because ‘ $E(x)p$ ’ that ‘ $E(x)E(x)p$ ’). By contrast, even if ‘ $E(x)p$ ’ might admit of some kind of explanation, it is presumably basic in the sense that it is not explained in further essentialist terms—and, for that matter, most certainly not in terms of an essence that derives from it.

---

<sup>45</sup> Dasgupta (2014: 591) holds that claims of consequential essence are grounded in claims of constitutive essence. A similar reasoning applies to the so called ‘truth-regress’.

<sup>46</sup> Moreover, if, say, God obliterated ‘ $E(x)p$ ’, I contend that the collapse of all the iterated essences above ‘ $E(x)p$ ’ in **S1** would necessarily ensue.



As soon as the idea that the essences in **S1** constitute an upwards chain of explanations is fully appreciated (*i.e.*, a chain that proceeds from ‘ $E(x)p$ ’ to ‘ $E(x)E(x)p$ ’, etc.), options **( $\alpha$ )** and **( $\beta$ )** break down. So even if, so far, the question of what explains ‘ $E(x)p$ ’ is left unanswered, the iterated essences in **S1** admit of an explanation. Consequently, **( $\alpha$ )** is false. And since the direction of explanations among essences proceeds upwards, it can hardly be maintained that **( $\beta$ )** is true on pain of vicious circularity: since ‘ $E(x)p$ ’ explains ‘ $E(x)E(x)p$ ’ there would be, in turn, an explanation of the former claim in terms of the latter, which violates both the asymmetry and irreflexivity of explanation.<sup>47</sup> (Also, a constitutive essence would be both prior and posterior to itself, as well as dependent on something that depends upon it. This is untenable.) Furthermore, such a downwards chain of explanations would also go against the plausible idea that explanations flow from the simpler facts to the more complex ones, and that there are no infinitely descending chains of explanation.<sup>48</sup> The way essences are structured in **S1** has important consequences for the thesis according to which essence is the source of all metaphysical necessities, as I will now explain.

Let us consider ‘ $E(x)p$ ’: this essence does not only provide an explanation of ‘ $\Box p$ ’ but also of ‘ $E(x)E(x)p$ ’. In accordance with the strategy considered so far, this latter iterated essence is needed to explain ‘ $\Box E(x)p$ ’ (which, again, is directly obtained since **NE** is true of ‘ $E(x)p$ ’) and to uphold **Source**. Hence, we arrive at the

---

<sup>47</sup> Worries in connection to option **( $\beta$ )** seem to find their source in a principle advocated by Rosen (2010: 119): if  $E(x)p$ , then  $[p]$  is grounded in  $[E(x)p]$ . However, if true, this principle can presumably not be applied unrestrictedly. For, that would be viciously regressive:  $[E(x)p]$  would depend on infinitely many essence-facts. See also Dasgupta (2014) and Van Cleve (2018) for a discussion and related difficulties that Rosen’s principle triggers.

<sup>48</sup> *Cf.*, Rabin and Rabern (2016) and Van Cleve (2018) for similar points.

following schematic instance of how explanations of the necessity of essences are structured in **S1**:

$[E(x)p] < [E(x)E(x)p]$ ,  $[E(x)E(x)p] < [\Box E(x)p]$ ;

etc.

I find such explanations suspicious, because ‘ $E(x)E(x)p$ ’ is explained by ‘ $E(x)p$ ’ and, in turn, ‘ $E(x)E(x)p$ ’ explains something that pertains to ‘ $E(x)p$ ’: its modal status. I think that essentialists are getting dangerously close to a circularity. And, in fact, there is only one small step to take to reach such a result.

The reason is that, quite independently from what Fine’s theory of essence has to say about explanations of necessity *per se*, the connection between a given essence and its prejacent proposition is considered to be an explanatory one.<sup>49</sup> That is, if ‘ $E(x)p$ ’, then ‘ $p$ ’ is explained by ‘ $E(x)p$ ’. To illustrate, Socrates is human because Socrates is essentially human.<sup>50</sup> Undoubtedly, such explanations are metaphysical in nature, and when combined with **Source**, the resulting picture is one where essence does not only support explanations of why a given proposition,  $p$ , is (metaphysically) necessary, but also of why  $p$  is true: there is no ‘division of labour’ between essence and some other phenomenon. I contend that this is

---

<sup>49</sup> As Glazier puts it, “it can hardly be denied that ‘ $t$  is essentially such that  $A$ ’ provides a metaphysical kind of explanation of  $A$ ” (2017: 2874). However, he disagrees with Rosen (2010), who flirts with the idea that the kind of metaphysical explanation at stake is grounding.

<sup>50</sup> The focus of the question is on Socrates himself. If the question targets being human and why it is predicated of Socrates rather than being a water molecule, then the *explanans* is presumably not Socrates’s essence. Thanks to Thomas Sattig for bringing my attention to this crucial distinction.

precisely how such explanations are supposed to go. For it would be strange to think that some phenomenon  $\Phi$  is explanatory relevant to  $p$ 's (metaphysical) necessity but not to  $p$ 's truth. Van Cleve makes a similar point:

If we say with Descartes that the Pythagorean Theorem is necessary only because it was established by God, we should also say that it is true only because it was established by God. To believe otherwise is to attribute to God a queer form of omnipotence that holds sway over truths of the form 'p' but not over truths generally. And similarly in other cases. For example, proponents of the linguistic theory of logical necessity should say (and did say) that necessary truths are true in virtue of meanings. Returning to Kant, if we say that the propositions of geometry owe their necessity to our cognitive constitution, we should also say that they owe their truth to our constitution. (1999: 40)

The rationale behind Van Cleve's examples is that whatever makes something necessary also makes it true. These observations can be generalized so as to obtain the following principle that I find plausible:

**NT** : If  $[A] < [\Box B]$ , then  $[A] < [B]$ .<sup>51</sup>

I contend that **NT** holds only for metaphysical necessity and what stands for its source. I take no particular stance on whether **NT** is true for other kinds of

---

<sup>51</sup> The formulation of this principle is inspired by several remarks made by Van Cleve in his (1999) and (2018).

necessity and their respective source(s), such as, for instance, normative necessity. Thus, applied to the essentialist framework, **NT** amounts to saying that if an essence explains why a proposition  $p$  is necessary, then that same essence explains why  $p$  is true. And this should be no surprise: after all, given that essentialists endorse **Source**, they should have no reason to deny that, given a particular necessity, essences fit the role of the *explanans* in **NT**'s antecedent. And, to my knowledge, no essentialist denies the explanatory link between a constitutive essence and its prejacent.<sup>52</sup> Thus, it should pose no problem for essentialists to say that essence plays the role of the *explanans* in the consequent of **NT** too.<sup>53</sup> To illustrate what has just been said, consider the following essentialist instance of **NT**: if  $[E(x)p] < [\Box p]$ , then  $[E(x)p] < [p]$ . Note, however, that it may well be the case that, depending on what stands for 'A' in **NT**, A metaphysically explains B together with some law. For simplicity, I will assume here that if A is substituted for essence in **NT**, then A fully explains B. This should not substantially affect the following arguments.

Now, **NT** brings problematic consequences onto **S1** and raises several important questions for the metaphysics of essence. One of the main difficulties is that since  $[E(x)p]$  explains  $[E(x)E(x)p]$  and since, in turn,  $[E(x)E(x)p]$  explains  $[\Box E(x)p]$ , it follows by **NT** that  $[E(x)E(x)p]$  should also explain  $[E(x)p]$ . This constitutes a

---

<sup>52</sup> Cf. Glazier (2017), Rosen (2010), and Zylstra (2019).

<sup>53</sup> One might think that **NT** holds solely on the grounds that it is backed up by transitivity, via an explanation of B in terms of  $\Box B$ —*i.e.*, if  $[A] < [\Box B]$  &  $[\Box B] < [B]$ , then  $[A] < [B]$ . This might be true if what stands for 'A' is a phenomenon other than essence and if the necessity at stake is not metaphysical (or absolute). In the context of essence, however, even though I do not deny that ' $\Box B$ ' entails 'B', I contend that **NT** holds without the help of transitivity. This is because of the intimate connection between an essentialist truth and its prejacent: it is because it is true in virtue of the nature of  $x$  that  $p$  that  $p$  is true, and not because  $p$  is necessary that  $p$  is true. The latter is merely a 'consequence' of the view.

direct violation of the asymmetry of explanation—*i.e.*,  $[E(x)p] < [E(x)E(x)p]$  and  $[E(x)E(x)p] < [E(x)p]$ . Also, with the assumption that metaphysical explanation is transitive, it follows that  $[E(x)p]$  eventually explains itself, which violates the irreflexivity of explanation. In brief, **S1** is structurally problematic since it involves circular explanations.<sup>54</sup> So something has got to give. As explained above, we have no reason to think that, in **S1**, an iterated essence explains a constitutive essence (or, perhaps, an essence being ‘simpler’ than itself (*cf.* Van Cleve (2018))). Consequently, since it cannot be the case that  $[E(x)E(x)p]$  explains  $[E(x)p]$ , then, by *modus tollens*,  $[E(x)E(x)p]$  does not explain  $[\Box E(x)p]$ .<sup>55</sup>

If what has been said so far is true, then we no longer have a structure of essences and necessities like **S1**, but rather a different explanatory schemata that takes the following form:

2.  $[\Box E(x)p]$ ;
1.  $[E(x)p] < [\Box p]$ .

On this view, the necessity of ‘ $E(x)p$ ’ lacks an explanation in terms of essence and **Source** is false (since not all necessities find their source in essence). In fact, **NT** neatly establishes that if ‘ $E(x)p$ ’ lacks an explanation (in terms of essence), then so does its necessity (more on this below). My point, then, is that since we cannot appeal to iterated essences (with bearer identity) to explain the truth of constitutive

---

<sup>54</sup> Difficulties also arise in terms of priority and dependence. Eventually, ‘ $E(x)E(x)p$ ’ depends upon ‘ $E(x)p$ ’, but ‘ $E(x)p$ ’ also depends upon ‘ $E(x)E(x)p$ ’ since **NT** holds. Each fact is therefore prior to the other and, by transitivity, dependent upon and prior to itself.

<sup>55</sup> Perhaps we should be cautious about what kind of metaphysical explanation is at stake in these arguments. I will return to this point shortly in the next section.

essences, then, by the same token, I claim that we cannot explain their necessity in that way. Given that the appeal to an iterated essence (with bearer identity) requires that such essences derive from and—consequently—are explained by a (corresponding) constitutive essence, there cannot be an explanation of the latter in terms of the former.

### 3.III Potential solutions on behalf of essentialists

So far, my arguments have relied on **Source**, **NE**, and **NT**, together with the plausible assumption that ‘ $E(x)E(x)p$ ’ is metaphysically explained by ‘ $E(x)p$ ’. In order to block the violations of asymmetry and irreflexivity identified in the previous section, essentialists can reject and/or modify some of these claims. But how? And which ones actually qualify? **Source** should not be denied because doing so would amount to weakening the explanatory power of the essentialist theory. Of course, one could impose a restriction on **Source** together with a disjunctive condition on explanations of necessity, but I will only appeal to such a strategy—and, more generally, to any kind of modification of **Source**—should it happen that no other option is available.<sup>56</sup> With respect to **NE**, essentialists could reject it on the grounds that essences are purely amodal (*i.e.*, neither necessary nor contingent). But here I must side with Wildman who explains that such a strategy commits one to modal gaps (2018: 8): how can something entirely devoid of any modal status

---

<sup>56</sup> Hale (2013) for instance argues that **Source** should be restricted to necessities that do not prefix an essentialist operator, as explained in chapter 2.

give rise to necessity?<sup>57</sup> Hence, the difficulties raised by the authors I have mentioned in chapter 2 will arise if the amodal move is endorsed. Finally, as explained earlier, we have good reasons to uphold the assumption according to which ‘ $E(x)p$ ’ provides a metaphysical explanation of ‘ $E(x)E(x)p$ ’, and not the other way around. By elimination, then, we are left with **NT**.

The first thing that comes to mind is that essentialists could simply reject **NT** on the grounds that since it triggers the substantial difficulties I have raised, it is a sign that it must be false (or, at least, that it does not hold for the essentialist framework). Here, however, I must disagree. For **NT** does not compromise the explanatory role of a constitutive essence towards the necessity and the truth of its prejacent: if  $[E(x)p] < [\Box p]$ , then  $[E(x)p] < [p]$ . Instances of such schematic explanations seems perfectly acceptable to me.

The best shot available to essentialists, then, is trying either to clarify how **NT** can be construed or to modify it, and I will now provide two examples of how this can be done. For the first kind of strategy, one could argue that my views on ‘metaphysical explanation’ are not fine-grained enough, because identifying which specific kinds of metaphysical explanations are involved in **S1** might be relevant to blocking my arguments. In particular, essentialists could try to argue that the explanatory relation at stake in the consequent of **NT** is substantially different from the explanatory relation that holds between constitutive essences and iterated essences. This would amount to arguing that even if, say, ‘ $E(x)p$ ’  $\phi$ -explains ‘ $E(x)E(x)p$ ’ and ‘ $E(x)E(x)p$ ’  $\psi$ -explains ‘ $E(x)p$ ’, no violations of asymmetry and

---

<sup>57</sup> See also Mackie (2020).

irreflexivity ensue because  $\phi$ -explanations and  $\psi$ -explanations cannot be chained together.

However, this strategy faces several challenges. To see why, essentialists have to identify what are the specific kinds of metaphysical explanations that can be substituted for  $\phi$  and  $\psi$ . With respect to the former, grounding seems to be the only viable option, which means that ‘ $E(x)p$ ’ is grounded in ‘ $E(x)$ ’.<sup>58</sup> And with respect to  $\psi$ , there are not many candidates to choose from. For obvious reasons, we cannot appeal to grounding and so we have to find another kind of metaphysical explanation that can govern essences and their preajacent. To my knowledge, the only candidate in the literature is Glazier’s essentialist explanation—a *sui generis* kind of metaphysical explanation that takes the following form: “A because t is essentially such that A” (2017: 2873).<sup>59</sup> The question for essentialists is whether Glazier’s essentialist explanations can be chained with grounding. Yet, regardless of a decisive answer on that matter, I maintain that it will be of no help. For, as Glazier explains, essentialist explanations are only meant to govern the constitutive notion of essence, and they are ultimate—*i.e.*, their *explanantia* do not admit of an essentialist explanation. This means that if essentialists construe the consequent of **NT** as involving essentialist explanation, they can only use this principle at the level of ‘ $E(x)p$ ’ in **S1**. So, since ‘ $E(x)p$ ’ essentially explains ‘ $p$ ’, it follows by ultimacy that ‘ $E(x)p$ ’ cannot be essentially explained by ‘ $E(x)E(x)p$ ’. Therefore, if

---

<sup>58</sup> Cf. Dasgupta (2014).

<sup>59</sup> See Glazier (2017: §2) for arguments to the effect that essentialist explanation is *sui generis* and that, therefore, it should not to be conflated with grounding (and other kinds of metaphysical explanations).



' $E(x)E(x)p$ ' does not explain ' $E(x)p$ ', then, by *modus tollens* on this essentialist construal of **NT**, ' $E(x)E(x)p$ ' does not explain ' $\Box E(x)p$ '.

Importantly, the success of this strategy depends entirely on there being a specific kind of metaphysical explanation that does not only hold between essences and their preajacent, but that can also not be chained with grounding.<sup>60</sup> In the absence of such a candidate-explanation that can be substituted for  $\psi$ , I contend that this strategy is useless for blocking the violations of asymmetry and irreflexivity in **S1**.

The second kind of strategy essentialists can opt for consists in modifying **NT**. One way to do so is as follows:

**NT\***: if  $[A < \Box B] \ \& \ \neg [B < A]$ , then  $[A < B]$ .

With **NT\***, essentialists can uphold the idea that iterated essences explain the necessity of other essences.<sup>61</sup> That is, since the antecedent of **NT\*** will always be false if ' $A$ ' and ' $B$ ' stand for essence-truths, inferences to instances of ' $A < B$ ' are blocked. However, I fear that such a modification rests on purely *ad hoc* grounds: one would endorse it solely to avoid difficulties that arise at the level of iterated essences in **S1**. And even if this were not the case, I am inclined to think that whatever the source of necessity is, it should provide explanations uniformly and in

---

<sup>60</sup> One might disagree with my claim that grounding is the relation at stake between ' $E(x)p$ ' and ' $E(x)E(x)p$ '. To be fair, I do not know what other kind of explanation can be appealed to in this case. But either way, this explanation,  $\phi$ , must be such that it cannot be chained with whatever explanation is substituted for  $\psi$ . In the absence of a better candidate than grounding, I think that my point holds.

<sup>61</sup> Thanks to Tobias Wilsch for this suggestion.

accordance with **NT**—especially if this principle works just fine with constitutive essences and their prejacent. So, to say that ‘ $E(x)E(x)p$ ’ is only relevant to the necessity of ‘ $E(x)p$ ’ strikes me as unwarranted.

Further investigation is required to figure out whether essentialists can bring other kinds modifications upon **NT** in order to block the violations of asymmetry and irreflexivity. Here, I merely outlined two strategies and explained why, in my opinion, they fail. I will leave it at that for the time being.

Thus, what emerges from this discussion is that there is a schism within the explanatory framework of essence: on the one hand, **NT** works just fine with constitutive essences and their prejacent, but, on the other hand, it fails to deliver the expected results at the level of iterated essences (with bearer identity). What kind of conclusion should we draw from this? I contend that just because **NT** fails at the level of iterated essences (with bearer identity), it is not sufficient to conclude that **NT** is false. Rather, I think that the proper conclusion to draw is that the necessity of a given essence should not be explained by a fact that derives from that very same essence.

So far, I have investigated the idea that what explains the necessity of ‘ $E(x)p$ ’ is the essence of some  $y$ , such that  $y = x$ . I have identified what kind of difficulties follow from this assumption, and I will now continue my investigation with the assumption that  $y \neq x$ .

### 3.IV Iterated essences without bearer identity

In this section, I will try to explain how the necessity (and also the truth) of ‘ $E(x)p$ ’ might be accounted for if appeal to iterated essences without bearer identity is made. This strategy consists in saying that ‘ $E(x)p$ ’ owes its necessity to the *constitutive* essence of some  $y$ , such that  $y \neq x$ : ‘ $\Box E(x)p$ ’ because ‘ $E(y)E(x)p$ ’.<sup>62</sup> But what can  $y$  be? It is difficult to find an answer. In his (2017), Glazier argues that there is no entity  $y$  such that it is true in virtue of the constitutive essence of  $y$  that ‘ $E(x)p$ ’. Even though Glazier discusses this point to sustain the claim that essentialist explanations are ultimate, I think that it is an important question that essentialists ought to have in mind. Generally speaking, asking whether constitutive essences belong to the essence of some  $y$  amounts to asking whether such essences can be explained.

Glazier mentions that  $y$  could be Essence itself. (Hereafter, I use the capital letter ‘ $E$ ’ to refer to essence *qua* essence.) This would mean that constitutive essences of the form ‘ $E(x)p$ ’—such that  $x \neq$  Essence—are essential to Essence. But a few problems immediately arise for this view. All constitutive essence-claims will be essential to Essence: the facts that Socrates is essentially human, that Plato is essentially human, etc. The difficulty is that it does not seem that we are in the business of characterizing what Essence is if we say that Essence is such that Socrates/Plato is essentially human. Intuition suggests that Essence knows nothing of the particular other essences (and, by the same token, of the entities they are

---

<sup>62</sup> In fact, it might be the case that only constitutive essences can give rise to necessity, but I am not sure whether this can be conclusively established.

about). It seems clear that propositions that are essential to Socrates do not characterize what the nature of Essence is. Thus, appeal to the essence of Essence is presumably at odds with the plausible idea that, as Glazier explains, “‘the ‘essentialist source’ of a proposition, that in whose nature the proposition lies, must itself be a (Russellian) constituent of the proposition” (2017: 2887). Technically, this amounts to saying that, in the proposition ‘E(Essence)E(x)p’, Essence must be a constituent of ‘E(x)p’. Glazier thinks that since Essence is—by assumption—distinct from  $x$  and not a constituent of ‘ $p$ ’ (because ‘ $p$ ’ is a truth about  $x$ ), ‘E(x)p’ is the claim that “some proposition not involving [E]ssence lies in the nature of something other than [E]ssence” (*Ibid.*). So the question remains: can ‘E(x)p’ be true in virtue of the essence of Essence? Glazier admits that he has no proof against this claim, but that, to his knowledge, there is no plausible example of this sort.

Perhaps, there is a way to make sense of the essence of Essence proposal, but I will not try to give a definitive answer on that matter here. Rather, for the sake of the argument I will assume that both ‘E(x)p’ and its necessity can be explained by the constitutive essence of some entity,  $\pi$ , in order to see whether this strategy can be of any help. Thus, if you are happy with Essence, then just read the following arguments as involving Essence.<sup>63</sup>

---

<sup>63</sup> Other candidate entities can be God or, one idea I am sympathetic to, Leibniz’s best of all possible worlds. If we can truly say that the best of all possible worlds is essentially such that Socrates is human, that Plato is human, etc., then this could be used as a counterexample to the constituency and ultimacy constraints highlighted by Glazier for his essentialist explanation. However, we should make sure that when we say that the best of all possible worlds is essentially such that Socrates is human, it is the constitutive notion of essence that is at stake, and not some other essentialist notion.

Consider ‘ $E(\pi)E(x)p$ ’: given **NE**, it must be necessary. In order to uphold **Source**, essentialists face two possibilities: either ‘ $\Box E(\pi)E(x)p$ ’ is explained by an iterated essence with bearer identity or by an iterated essence without bearer identity—*i.e.*, either by ‘ $E(\pi)E(\pi)E(x)p$ ’ or ‘ $E(z)E(\pi)E(x)p$ ’. Opting for the latter forces essentialists to investigate further away from  $x$  and  $\pi$ , and I have honestly no idea what  $z$  could even stand for at this point. Other things being equal, an infinite regress ensues because the question now switches to which entity is such that facts about its essence can explain the necessity of ‘ $E(z)E(\pi)E(x)p$ ’, and so on. With respect to the former option, essentialists unsurprisingly arrive at a structure similar to **S1**:

**Structure 2 (S2)**

(h):  $[E(\pi)\dots E(x)p] < (g) [\Box E(\pi)\dots E(x)p]$ ;

(f):  $[E(\pi)E(\pi)E(\pi)E(x)p] < (e) [\Box E(\pi)E(\pi)E(x)p]$ ;

(d):  $[E(\pi)E(\pi)E(x)p] < (c) [\Box E(\pi)E(x)p]$ ;

(b):  $[E(\pi)E(x)p] < (a) [\Box E(x)p], (a') [E(x)p]$ .

In **S2**, there is an explanation of both ‘ $\Box E(x)p$ ’ and ‘ $E(x)p$ ’ in terms of ‘ $E(\pi)E(x)p$ ’: (a) and (a’) are explained by (b). Such schematic explanations can be considered as unproblematic since they are arguably valid instances of **NT**: if  $[E(\pi)E(x)p] < [\Box E(x)p]$ , then  $[E(\pi)E(x)p] < [E(x)p]$ . As a matter of fact, these explanations mirror

how ‘ $E(x)p$ ’ explains both the truth and the necessity of its prejacent in **S1**: ‘ $\Box p$ ’ and ‘ $p$ ’ because ‘ $E(x)p$ ’.<sup>64</sup>

However, difficulties arise at the level of ‘ $E(\pi)E(x)p$ ’ and here essentialists face another two options. The first one is to say that, just as **S1**, **S2** is composed of one constitutive essence, ‘ $E(\pi)E(x)p$ ’, and an infinity of iterated consequential essences above it—*i.e.*, ‘ $E(\pi)(E(\pi)E(x)p)$ ’, etc. If this is the option endorsed, then the exact same structural problems as the ones encountered with **S1** resurface. This is so because, given **NT**, violations of asymmetry and irreflexivity ensue:  $[E(\pi)E(x)p] < [E(\pi)(E(\pi)E(x)p)]$ ,  $[E(\pi)(E(\pi)E(x)p)] < \Box[E(\pi)E(x)p]$ , and then, by **NT**,  $[E(\pi)(E(\pi)E(x)p)] < [E(\pi)E(x)p]$ , contradiction.

The second option is to confer the same metaphysical status to all claims about the essence of  $\pi$  (*i.e.*, all claims of the form ‘ $E(\pi)E\dots$ ’), and to say that they are of the constitutive kind. Perhaps this is more plausible than saying that iterated essences in **S1** are constitutive, because truths about the essence of  $\pi$ —depending of what ‘ $\pi$ ’ actually refers to—might be said to characterize directly what  $\pi$  is, and nothing else. I acknowledge that this becomes a rather thorny matter, and we sink into ever deeper levels. But, to cut the discussion short, we saw what kind of problems arise if there is an upwards or downwards chain of explanations in such structures. Presumably then, no truth about the essence of  $\pi$  would be explained in terms of the essence of  $\pi$  if this second option is endorsed. This means that there can be no true instance of **NT**, because if  $[E(\pi)E(\pi)E(x)p] < [\Box E(\pi)E(x)p]$ , then **NT** tells us

---

<sup>64</sup> Note, however, that Glazier’s essentialist explanation cannot hold in this specific context since if ‘ $E(x)p$ ’ explains ‘ $p$ ’, there can be no essentialist explanation of ‘ $E(x)p$ ’ in terms of ‘ $E(\pi)E(x)p$ ’ (due to his ultimacy constraint).

that  $[E(\pi)E(\pi)E(x)p] < [E(\pi)E(x)p]$ . However, with what has just been said, it is not the case that  $[E(\pi)E(\pi)E(x)p] < [E(\pi)E(x)p]$ . Therefore, by *modus tollens*, it is not the case that  $[E(\pi)E(\pi)E(x)p] < [\Box(E(\pi)E(x)p)]$ .

So, explaining the necessity of ‘ $E(x)p$ ’ in terms of ‘ $E(\pi)E(x)p$ ’ can only partially solve the problems that essentialists face. For they now have to explain the necessity of ‘ $E(\pi)E(x)p$ ’ and, even if we assume that there is a candidate-entity that stands for ‘ $\pi$ ’, none of the options I have outlined in this section seems promising or successful.

Our discussion so far shows that there is no straightforward way for essentialists to account for the necessity of essences. On the one hand, if they appeal to an iterated essence with bearer identity to explain the necessity of ‘ $E(x)p$ ’, a structure like **S1** arises. *Prima facie*, such a structure is benign, but once plausible assumptions and principles are taken on board, violations of asymmetry and irreflexivity ensue, and there is no obvious way out. If, on the other hand, essentialists appeal to an iterated essence without bearer identity to explain the necessity of ‘ $E(x)p$ ’, a structure like **S2** is generated and their first challenge is to find which entity—if any—is such that it is essential to it that ‘ $E(x)p$ ’. The problem is that even if we assume, for the sake of the argument, that *there is* such an entity, essentialists merely transfer the initial difficulty of explaining ‘ $\Box E(x)p$ ’ to that of explaining ‘ $\Box E(y)E(x)p$ ’: either the *explanans* is an iterated essence with bearer identity or it is an iterated essence without bearer identity. The former option leads to vicious circularities while the latter forces essentialist to find another entity, thereby moving the same dilemma to a further level, and so on. Since there is no straightforward way out of these issues,

it seems to me that **Source** should be restricted to constitutive essences only, and that a different account of the necessity of such essences is needed.

This might be done by saying that essences explain their own necessity. Of course, this means that **NT** can no longer hold, because otherwise essences will eventually explain themselves: if  $[E(x)p] < [\Box E(x)p]$ , then  $[E(x)p] < [E(x)p]$ .<sup>65</sup> In response, essentialists could try to argue that if an essence explains its necessity, then that essence does not need to make itself true, because its truth is already given/established. However, the difficulties that pertain to gap will have to be considered: if it is indeed true that essences *simpliciter* lack any explanatory power to bring forth necessity, then on this view essences would not only lack the power to explain their own necessity, but also the power to explain other necessities. We could dig further into these issues but so far the prospects seem dim, and I rest my case regarding the gap and the necessity of essence. In the next chapter, I want to focus on a different framework, one where essences are not understood as real definitions, to see whether we can overcome these two core issues: the gap and the necessity of essence.

---

<sup>65</sup> Similarly, **NT\*** will not do: if  $[E(x)p] < [\Box E(x)p] \ \& \ \neg[E(x)p] < [E(x)p]$ , then  $[E(x)p] < [E(x)p]$ . Contradiction.



## Chapter 4: Essence and generalized identity

In the present chapter, I will set aside the idea that essences are construed as real definitions, and I will appeal to the notion of generalized identity to see whether the two difficulties faced by essentialists—namely, the explanatory gap challenge, and explaining the necessity of essences—can be overcome. The plan is the following: in section 4.I, I introduce the notion of generalized identity, spell out its properties and explain how it is similar to—and yet importantly different from—*objectual identity*. In section 4.II, I articulate the idea that essentialists can defend an entailment from essence to generalized identity. Section 4.III details that entailment-strategy: I explain how the (alleged) gap between essence and modality can be bridged, and why I think that the outcome is not compelling. In section 4.IV, I turn to Fabrice Correia and Alexander Skiles’s essentialist framework and I explain how they intend to address the explanatory-gap challenge, and whether they can explain the necessity of essences. I then end this chapter by giving my own conclusions on essence, generalized identity, and necessity.

### 4.I Generalized identity

In both philosophy and classical logic, the canonical way to construe the notion of identity is as a *relation* that every object bears to itself and to nothing else. Identity

is primarily thought of and used as an *objectual* relation in debates and arguments.<sup>66</sup> To that effect, Linnebo observes that, “[a]t the level of syntax, this is reflected in the fact that the identity predicate can meaningfully be flanked only by singular terms, not by predicates, complete sentences, or any other type of expression” (2014: 466). But recently, Rayo (2013) has shed light on a category of statements that he labels *just-is statements*, and which are also known as *generalized identities* in the literature (*cf.* Linnebo 2014, Correia & Skiles 2019). One might feel unfamiliar with these technical terms, but probably not with what they actually stand for, regardless of one’s philosophical background, as the following examples tend to illustrate:

- (a) For Socrates to be in pain *just is* for Socrates to be in a state of C-fibers firing;
- (b) For  $x$  to know that  $p$  *just is* for  $x$  to have a justified true belief that  $p$  (Correia & Skiles 2019);
- (c) For something to be a vixen *just is* for it to be a female fox (Hale 2013);
- (d) For an action to be good *just is* for it to be conform to the categorical imperative;
- (e) For  $F$  to be an essential property of  $x$  *just is* for  $F$  to be a necessary property of  $x$ ;
- (f) “For the number of the Dinosaurs to be Zero just is for there to be no dinosaurs.” (Rayo 2013: 3)

---

<sup>66</sup> See Kripke’s arguments in chapter 2.

Some of these examples—especially (a), (d) - (f)—are contentious.<sup>67</sup> My point, however, is not to take a particular stance towards these claims nor is it to defend their truth: they involve further arguments and commitments to specific theories to be successfully defended; a dualist about the mind and body will deny (a), just as a consequentialist and a Finean essentialist will (and should) reject (d) and (e), respectively. Rather, I merely use these examples to illustrate *how* the notion of generalized identity is put into work in the recent literature.

In all these examples the ‘is’ expression from ‘just is’ does not behave like a copula, but rather in the same way as the objectual identity relation does. That is, just as Hesperus *is* (the same *thing* as) Phosphorus, namely the planet Venus, the idea behind a given generalized identity statement, say (c), is that there is *no difference* between *being a female fox* and *being a vixen*; or, similarly with (a), that what it takes for an action to be good *is the same as* what it takes for that action to conform to the categorical imperative. These sentences are meant to describe a particular phenomenon or ‘portion of reality’, but under a different ‘guise’.

Despite this specific similarity, there are numerous aspects under which the relations of objectual and generalized identity resemble and differ ontologically. To begin with, as opposed to (rigid) designators that flank each side of the objectual identity relation, phrases like ‘for an action to be good’ do not seem to be the kind of linguistic expression whose task is to *refer* to something (that exists). To that effect, we should be careful not to think that generalized identities merely consist

---

<sup>67</sup> Cf. Rayo (2013: 4-5) for additional examples and explanations.

in a sophisticated way to talk about identity between *properties*. To illustrate this point, consider the following example put forward by Cian Dorr (2016: 40-1):

(g) To be a non-self-instantiator *just is* to fail to instantiate oneself.

If (g) were merely a property-identity statement in disguise, it could be assimilated to:

(g') The property of *being a non-self-instantiator* is the property of *failing to instantiate oneself*.

But as Bertrand Russell (2020) has pointed out, there is no such thing as the property of *being a non-self-instantiator*. Accordingly, the expressions flanking 'is' in (g) and (g') belong to different syntactic categories (Dorr 2016). By parity of reasoning, we would not say that the sentence 'I hope to be a good husband' is about a property, for if it were, it would be equivalent to 'I hope the property of being a good husband' (*ibid.*). The latter claim is meaningless, and the idea is that we ought to refrain from construing generalized identity statements as an ontologically committing way to talk about properties. For quite similar reasons, we should not construe generalized identity claims as identities between *facts* or *propositions*.

So what kind of theoretical purpose is generalized identity supposed to achieve? Roughly, a given generalized identity claim is meant to capture an identity between two (different) descriptions of one and the same portion of reality. That is, a

generalized identity establishes a connection between two *sentences* that represent the same phenomenon ‘out there’ in the world. However, in the examples provided so far, the ‘just is’ expression is not, *prima facie*, flanked by sentences—*e.g.*, syntactically speaking, neither ‘for an action to be good’ nor ‘for something to be a vixen’ are sentences. But this is merely a syntactical detail that can be ironed out within a proper formalization. To this effect, I will follow Rayo who uses a ‘just-is’ operator,  $\equiv$ , that can be indexed by zero or more variables, and that takes (open or closed) formulas in place of its arguments to yield another formula. To illustrate this point, examples (a) and (b) can be rewritten as follows, respectively:

(a’) Socrates is in pain  $\equiv_y$  Socrates is in a state of C-fibers firing.

(b’)  $x$  knows that  $p \equiv_x x$  has a justified true belief that  $p$ .

Moreover, I will rely on the following conventions to formalize generalized identity statements (*cf.* Correia & Skiles 2019: 5):

(1)  $p \equiv q$  (read: ‘for it to be the case that  $p$  just is for it to be the case that  $q$ ’);

(2)  $p \equiv_{x, y, \dots} q$  (read: for some things  $x, y, \dots$  to be such that  $p$  just is for these things to be such that  $q$ );

(3)  $Fx \equiv_x Gx$  (read: for  $x$  to be  $F$  just is for  $x$  to be  $G$ ).

Furthermore, the ‘just-is’-operator possesses the same relational properties as its objectual counterpart, namely: *reflexivity*, *symmetry*, and *transitivity*. This means, *inter alia*, that a generalized identity statement does *not* (and is not meant to) state

any order of *priority* between the sentences that flank each side of the ‘just is’-expression. Thus, ‘*x* is a female fox’ is no less or more fundamental than ‘*x* is a vixen’ or *vice versa*: both sentences provide a correct—and yet different—description of one and the same fact. In sum, sentences that flank each side of the just-is operator are, structurally speaking, at the same ‘metaphysical level’; one does not ground/metaphysically explain the other.<sup>68</sup> Consequently, the way generalized identity is construed here is not meant to achieve any definitional purpose—*e.g.*, in (e) an essential property of *x* is not *defined* in terms of a necessary property of *x* (and *vice versa*).<sup>69</sup> Rather, someone who endorses this claim accepts the thesis that it makes *no difference*, in the world ‘out there’, to instantiate a necessary property or an essential property. That being said, it is important to stress that there is a *conceptual* (and representational) difference at stake. To see why, consider (a) ‘Socrates is in a state of C-fibers firing’: this sentence, conceptually speaking, carves closer to reality’s joints than ‘Socrates is in pain’, because it takes scientific evidence to discover that there are states of C-fibers firing, whereas we naturally experience on our own what being in pain is at the earliest stages of our existence. Again, the point here is that both claims consist in different ways of describing the same fact.

---

<sup>68</sup> Let me stress that there is an unworldly/representational reading of these claims where it can meaningfully be said that one is prior to the other—*e.g.*, where ‘being a female fox’ is more fundamental than ‘being a vixen’. However, I am interested here in the worldly/non-representational reading of these claims and with how they describe reality.

<sup>69</sup> In his (2017), Correia uses the notion of generalized identity to give an account of real definitions (and hence to define different notions). However, this is importantly different from what is at stake here, because Correia imposes a constraint of metaphysical priority on his account.

With respect to truth-conditions, generalized identities are *true* on the condition that each sentence flanking the ‘just is’-expression *fully* and *accurately* describes the same feature of reality (Rayo 2013: 5). So what is required of reality for, say, (d) to be true is that there is no difference between someone performing a good action and someone performing an action conforming to the categorical imperative; the sentences flanking the ‘just’-is expression in (d) are indeed different, but they describe the same phenomenon. In other words, a generalized identity says that one and the same fact can be *truly* described in (at least two) different ways. To give an analogy, two cameras with different captors and *f*-ratio lenses can be used to take a picture of the *same* subject. When printed on paper, both pictures *do* depict the same subject, but differently: perhaps under a different light, angle, fineness of grain, contrast, etc. What matters, then, is that, say, your shot and my shot of the Boudhanath Stupa in Nepal depict—though differently—the same subject: these pictures are accurate shots of the same thing.

Next, like its objectual counterpart, the ‘just-is’-operator obeys to a certain version of *Leibniz’s Law*. In the objectual case, we know that, given two arbitrary rigid designators ‘*a*’ and ‘*b*’, if  $a = b$  and  $\phi a$ , *a* can be substituted with *b* in  $\phi$  (for some suitable  $\phi$ ) *salva veritate*. In the case of generalized identity, a similar law can be formulated for the sentential and predicational modes respectively:

**Leibniz’s Law for G.I.**

$$(p \equiv_v q) \supset (\phi \supset [q/p]\phi)$$

(Where  $[q//p]\phi$  results from sentence  $\phi$  by replacing one or more occurrences of  $p$  by  $q$ , with the condition that no variable that is free in  $p \equiv_v q$  is bound in  $\phi$  or  $[q//p]\phi$ .)

$(Fx \equiv_x Gx) \supset (\phi \supset [Gx//Fx]\phi)$

(Where  $[Gx//Fx]\phi$  results from sentence  $\phi$  by replacing one or more occurrences of  $Fx$  by  $Gx$ , with the condition that no variable that is free in  $[Fx \equiv_x Gx]$  is bound in  $\phi$  or  $[Gx//Fx]\phi$  (Dorr 2016: 48).)

For reasons similar to those that arise with objectual identity, this version of Leibniz’s Law cannot hold in *all* circumstances and, accordingly, it should be restricted. In particular, quotational and propositional attitude contexts are known to be *opaque*—*i.e.*, one cannot always substitute, in the objectual case, co-referring designators *salva veritate* within such contexts.<sup>70</sup> Typically, the objectual version of Leibniz’s Law is restricted to terms that are *directly* referring, such as ‘Socrates’, ‘Plato’, etc. But with respect to generalized identity, matters are somewhat trickier.<sup>71</sup> That being said, for the scope of the present inquiry, we can commit ourselves to exclude explicitly known cases of opaque contexts such as the ones mentioned above.

---

<sup>70</sup> For instance, if Sam knows that Bruce Wayne is a millionaire, it is not necessarily the case the Sam knows that Batman is a millionaire, even though Bruce Wayne is Batman. See chapter 1 as well.

<sup>71</sup> See Linnebo (2014) for a discussion.



Let us take stock. We have seen that the notions of objectual identity and generalized identity share important characteristics. At this stage, we may think that one is maybe more fundamental than the other in the sense that all discourse related to one notion can be reduced and explained in terms of the other. However, it does not seem possible to reduce generalized identity talk to objectual identity talk. The main point against this idea is that, as explained above, generalized identity is not in the ‘designating business’, while this seems to be one of the key-characteristics of objectual-identity.

What about the other way around? That is, can objectual-identity statements be captured in terms of generalized identity statements? This is a tricky question whose answer amounts to a different construal of objectual-identities. For instance, assume that  $a$  is identical to  $b$  ( $a = b$ ). We can translate this claim in the target-language as follows: ‘for something to be  $b$  just is for that thing to be  $a$ ’. But there are reasons to think that the grammatical surface of this generalized identity statement hides the *true* form of the expressions ‘to be  $b$ ’ and ‘to be  $a$ ’. Presumably, what is meant by these latter claims is ‘to be *identical* to  $a/b$ ’. To that effect, Adams (1981) and Plantinga (1983) both consider the predicates ‘being  $x$ ’ and ‘being identical with  $x$ ’ as coextensive—*i.e.*, as meaning the same thing. If this interpretation is correct, then it jeopardizes the possibility of construing objectual-identities in the language of generalized identity. For one thing, if ‘ $a = b$ ’ is translated as ‘ $\lambda x (x = b) \equiv_x \lambda x (a = x)$ ’, then we have not successfully reduced objectual identity to generalized identity since we make substantial appeal to ‘=’ in the new formulation. And, following Adams’s and Plantinga’s thoughts on the matter, we cannot rely on the claim according to which ‘to be  $x$  just is to be

identical to  $x$ ' for the same reasons. Consequently, it seems that either notion resists an analysis in terms of the other, and that we should construe them as two different species of identity relation belonging to what we may refer to as the 'identity' *genus*.

To summarize what has been said in this section so far, generalized identity is not to be construed as a special case of objectual identity and *vice versa*. The purpose of generalized identity claims does not consist in stating an identity between properties, facts, or even propositions, but rather to capture the idea that some particular phenomenon/fact can be truly and accurately described with two different descriptions: to say that for something to be  $F$  just is for it to be  $G$  is to say that, in reality, there is *no difference* between being  $F$  and being  $G$ .

## 4.II From essence to generalized identity

Let us now turn to generalized identity and essence. Correia and Skiles (2019) defend the idea that essence is analyzable in terms of generalized identity. Their framework is therefore different from Fine's: *inter alia*, the essentialist-operator is traded for the generalized identity operator, and questions about the nature of being  $x$  turn out to be questions about the identity conditions for being  $x$ . I doubt that many Finean essentialist will abandon their framework for Correia and Skiles's without good reason. But since they face the explanatory gap challenge and the problem of explaining the necessity of essences, this is perhaps where generalized identity can be put to work in improving the 'traditional' Finean framework (and its

reductive(non-reductive construals). Methodologically, rather than opting *directly* for Correia and Skiles's analysis of essence, I will start to discuss why these authors think that these notions are so intimately connected. Once this point is clarified, I will proceed to explain step-by-step how, in turn, the explanatory gap and the explanation of the necessity of essences might be overcome.

Our starting point takes the form of a question: 'why even think that generalized identity has any connection whatsoever with essence?'. One reason in favor of the idea that both notions are related is that they naturally (and equally) stop our demands for further (metaphysical) explanations. To illustrate, consider the following paradigmatic case with essences. One might want to ask '*why* is water composed of hydrogen and oxygen?', to which a plausible answer is that this is so *because* water is essentially composed of these chemical elements (or, if you prefer, because it lies in *the nature of* water to be so). With such an explanation, the fact that water contains both hydrogen and oxygen is explained in terms of the fact that it is of the essence of water to contain hydrogen and oxygen. But there is something puzzling if we try to inquire further into why water is *essentially* composed of H<sub>2</sub>O. Doing so would bring confusion to the table, and it is hard to know what to say except a mere reiteration of the *explanans* in slightly different terms, or to simply ignore the demand altogether. Dasgupta gives the following explanation of this phenomenon:

suppose [...] that it is essential to knowledge that someone knows only if she truly and justifiably believes. And suppose someone asks what explains this (in the metaphysical sense). In virtue of what (the question

is) is it part of what knowledge is that someone knows only if she truly and justifiably believes? It is difficult to know how to respond. One is tempted to say that this is just what knowledge is... but of course this is what we were asked to explain! In saying this one is most naturally heard not as trying to explain this fact about knowledge in any serious sense but rather as deflecting the demand for explanation. (2016: 386)

Echoing Dasgupta's observations, Glazier (2017) argues that when we have provided an explanation of *A* in terms of the fact that something is essentially such that *A*, the *explanans* admits of no further explanation in terms of essence. Of course, there might be another form of metaphysical explanation of essentialist facts, but it is hard to see what that might actually be.

The point, then, is that when one (successfully) gives an essentialist answer to a ('what') question, it seems that we have 'reached the end of the explanatory road' (Glazier 2017). Interestingly, the same is true of generalized identities. As Rayo explains, one key characteristic of these statements is that when we accept their *truth*, a *theoretical gap* is closed. That is, once semantical and/or cognitive demands for clarification are set aside, there is also something puzzling in asking the question 'why is it that *F*s are also *G*s?' if we accept the truth of 'for something to be *F* just is for it to be *G*'; the question strikes us as 'illegitimate' at best, and unintelligible at worst. To see why, consider what Rayo and Dorr say about demands for explanations of generalized identities:

Suppose [...] that someone says: ‘I can see that things composed of water are composed of H<sub>2</sub>O, but I wish to better understand why the world is such as to satisfy this condition.’ The natural reaction is to either find a charitable reinterpretation of the question (‘why are watery things composed of H<sub>2</sub>O?’) or reject it altogether (‘What do you mean *why*? For something to be composed of water *just is* for it to be composed of H<sub>2</sub>O’). (Rayo 2013: 55)

Identities are excellent *stopping places for explanation*; they do not cry out for explanation in their own right. Indeed, there is something odd about questions like ‘Why is Hesperus Phosphorus?’. [...] And this also applies to questions like ‘Why is it that to be a vixen is to be a female fox?’ Once we set aside the “remind me of reasons to believe” reading, and metalinguistic questions about the word ‘vixen’, it is hard to see what an answer would even look like. (Dorr 2016: 41, emphasis mine)

So to ask why either an essence or a generalized identity statement is true prompts the same kind of reaction; there seems to be no way to inquire further once that ‘level’ of truth is reached since they put a natural stop to our demands of explanation. Hence, there is a neat parallelism between essence and generalized identity with respect to (our demands for) metaphysical explanations.

Next, we saw in the first chapter that, prior to discussions about modality, the main goal of essentialist theories is to inquire into the nature of things, to figure out *what*

they *are*—*e.g.*, *what* Socrates and God *are*, respectively. Thus, (metaphysically loaded) ‘what’-questions bear the mark of essentialism and, accordingly, the essential characteristics of a given thing should point towards (at least) a (partial and) satisfactory answer to such questions. In a word, these characteristics are meant to specify *what it is* for a (kind of) thing to be that (kind of) thing. The main point I am going to make boils down to this: both essence and generalized identity statements can be considered as equivalent *answers* to ‘what’-questions. To the question ‘what is God?’, Finean essentialists will presumably answer that God is essentially an omnipotent, benevolent, and omniscient being. And with this kind of answer, explanations come to an end: it makes little sense to ask further *why* God is essentially omniscient.

And the situation is similar with generalized identity. When asking ‘*what* is God?’, it seems to be a perfectly acceptable answer to say that for something to be God just is for it to be omniscient, benevolent, and omnipotent: that is just *what it is* for something to be God. Furthermore, to ask *why* something that is omniscient, benevolent, and omnipotent also happens to be God is a confusing question, because for something to possess these characteristics altogether *just is* for that thing to be God.<sup>72</sup> It is therefore natural to think that both kinds of statements are intimately linked in some way given how they nicely answer to what-questions and put an end to our demands for further explanations. In a word, (specific) ‘why-questions’ come to an end with answers to (metaphysically-loaded) ‘what-

---

<sup>72</sup> Strictly speaking, there is a little mismatch between how essentialists and identity-theorists answer to what-questions. The way I have formulated these latter is object-oriented, and so are their answers in terms of essence. However, as we have seen generalized identity is not objectual. Thus, to the (objectual) question ‘what is God’ I have provided an answer which invokes the predicate expression to be God. I will return to this point shortly.

questions', and both essences and generalized identities seem to equally qualify in achieving this task.

Given these similarities, Correia and Skiles offer an account of essence in terms of generalized identity. This is an innovative way to think about essence and it constitutes the core of these authors's proposal. However, we do not need to embrace this framework yet. My purpose so far consists in improving Finean essentialism and not in abandoning it altogether for another theory. The idea I want to prioritize now is that essence *leads* to generalized identity by way of *entailment*. And after all, this should not be surprising: by providing answers to (metaphysically loaded) 'what'-questions, essences specify and determine the identity conditions of the entities they are about. Roughly, if it is essential to *a* to be *F*, it follows that for something, *x*, to be *a* just is for *x* to be *F*.<sup>73</sup> On this view, essence is 'richer' than generalized identity, since there can be generalized identity claims that do not derive from essences. For instance, the claim 'for the Camp Nou to be full just is for the Camp Nou to have 99354 spectators' is not derived from the essence of that stadium. For instance, it could have had a 99355 spectators capacity and this characteristic does not seem to pertain to its nature. Here, I do not defend the view according to which generalized identity is what determines the essence of things, but rather the other way around: the idea that every essence-claim entails a generalized identity claim.

---

<sup>73</sup> I ignore here matters about the partial vs. full conditions to be a thing. Arguably, there is more to being Socrates than just being human. I will return to this point later on.

As has been discussed, Finean essentialism (and also modalism, for that matter) is tailored around the *objectual* notion of essence, while generalized identity is, by definition, not objectual (since it does not take singular terms in place of its arguments). Hence, in order to defend the entailment from essence to generalized identity, we need to have a homogeneous language (if only for the sake of readability). In order to do so, we have to think outside of the ‘objectual box’ with the help of Correia’s notion of *generic essence* (2006).<sup>74</sup> In a word, generic essence is meant to characterize what Correia labels *ways of being*—*i.e.*, “the essential features of what predicates express” (*ibid.*: 2). So the theoretical purpose of generic essence is to answer to questions of the type ‘what is it to *F*?’ (where ‘*F*’ is a predicate expression such as ‘be a philosopher’) as opposed to ‘what is *x*?’ where the focus is on a specific entity.

As Correia explains, there is a natural inclination to think that “each generic statement is about the objectual nature of a property or of properties”, and therefore that we should construe “generic essentialist statements as objectual essentialist statements about [...] properties” (*ibid.*: 8). To give an example, this would mean that it is true in virtue of what the property of *being made of gold* is that whatever *is made of gold has atomic number 79*. But no matter how natural that inclination might be, Correia provides arguments to the effect that this view cannot be sustained. I will now briefly present two of them.

---

<sup>74</sup> Of course, this is not to say that generic essence was an unknown concept prior to Correia’s work, as readers of Aristotle should already be familiar with this notion. Rather, the point is that Finean essentialism has overlooked generic essence by thinking that it is merely a special case of objectual essence. Fine eventually acknowledges the distinction in his later works (2015).



First, talk of the objectual nature of properties ontologically commits one to properties, and one can have several reasons for rejecting such a commitment (*e.g.*, being a nominalist). In response, one might think that, since essentialists are already ontologically committed to concepts and logical concepts, it would not be absurd to believe in properties as well. Yet—and this is Correia’s *second* and main point—even if essentialists accept such a commitment, the problem is that statements of generic essence do not seem to be in the business of being ontologically committing at all. For, there seems to be nothing contradictory with the claim that round-squares are essentially round and the idea that there are no properties at all. And this, in my view, is Correia’s most important point, because proponents of such a view would have to defend “that despite appearances, generic essentialist statements do carry commitments to properties” (*ibid.*: 9). Given these remarks, I side here with Correia in saying that we should not opt for a designating reading of generic essence statements.

Now for Correia’s positive essentialist proposal. In his (2006) paper, he explains that just as Fine takes the notion of objectual essence as a primitive one, he is inclined to do the same with the notion of generic essence.<sup>75</sup> Correia opts for a specific framework of essence, whose tenet is that *every* essential truth can ultimately be spelled out in terms of generic essence. In order for this analysis to succeed, Correia proposes to associate a haecceity-predicate expression with each singular term—*e.g.*, the singular terms ‘Socrates’ and ‘Plato’ are associated with

---

<sup>75</sup> In their (2019) paper, Correia and Skiles defend the idea that generic essence is analyzable in terms of generalized identity. As stressed before, I will consider this option later on. For now, I am merely interested in a comprehensive way to cash out the entailment from essence to generalized identity.

the predicate expressions ‘*being Socrates*’ and ‘*being Plato*’ respectively. Then, the analysis of objectual essence in terms of generic essence proceeds as follows:

**Generic-Objectual Essence:** It is true in virtue of what *a* is that *p* iff it is true in virtue of what it is to be *a* (i.e., ‘*being a*’) that *p*.

Hence, Correia’s generic notion of essence is treated as the most basic notion of essence in his framework. An (objectual) essence-truth like ‘it is true in virtue of the nature of Socrates that Socrates is human’ can now be reformulated as ‘it is true in virtue of what it is to be Socrates that Socrates is human’. On this account, we have traded singular terms for predicate-expressions, and we have a unified account to talk about the essence of objects and ways of being. Since all essence-talk is ultimately expressible in terms of generic-essence, we can take a step further and explain how the connection between generic-essence and generalized identity can help essentialists address the difficulties they face.

#### 4.III The explanatory gap problem revisited

In this section, I will try to defend essentialism against the explanatory gap challenge that I have discussed in chapter 2. In order to do so, the central idea is to say that (Finean) essentialist claims *entail* generalized identity claims. That is, on this view the claim ‘it is essential to being *F* to be *G*’ entails the claim ‘to be *F* just is to be *G*’. The next step is to explain how generalized identity behaves with modality. As Kripke and Marcus have shown, there is a direct connection between

objectual identity and metaphysical necessity: if  $a$  is identical to  $b$ , then this is so as a matter of necessity; there is no world where  $a$  fails to be identical with  $b$ . With respect to generalized identity, there is also a straightforward path to metaphysical necessity. Here, I follow Correia and Skiles's (2019) conventions to illustrate this point:

**Predicational form**     if  $Fx \equiv_x Gx$ , then  $\Box \forall x \Box (Fx \text{ iff } Gx)$ ;

**Sentential form**        if  $p \equiv q$ , then  $\Box (p \text{ iff } q)$ .

Consider an example with the **predicational form**: if for  $x$  to be God just is for  $x$  to be omniscient, then, as a matter of necessity, for all  $x$ ,  $x$  is God if, and only if,  $x$  is omniscient; and if this is so, then we obtain the claim that necessarily  $x$  is God only if  $x$  is omniscient. Hence, generalized identity claims entail metaphysically necessary claims just as objectual identity does.

The gap between essence and necessity can be bridged as follows (using a generic essence claim together with Correia and Skiles's **predicational form** principle above) :

$E(F)G \rightarrow Fx \equiv_x Gx \rightarrow \Box \forall x \Box (Fx \text{ iff } Gx)$ ,

and consequently:  $\Box (Fx \supset Gx)$ .

So, against philosophers who argue that an explanatory gap is felt between essence and necessity, essentialists can answer that this gap is only illusory. For as soon as the entailment from essence to generalized identity is acknowledged, there is a

straightforward path to metaphysical necessity:<sup>76</sup> essence-claims entail generalized identities (*i.e.*, they entail the identity-conditions of what it is to be *F*, to *F*, etc.), and generalized identities lead to necessity—*i.e.*, the explanatory gap is bridged with Correia and Skiles’s **predicational/sentential form** principles. By way of entailment, then, essence leads to generalized identity which, in turn, gives rise to metaphysical necessity. One of the upshots is that it is (presumably) not (or, at least, less) controversial to say that generalized identity takes us to metaphysical necessity than it is to say that essence explains necessity.

Importantly, the entailment from, say, ‘*p* just is *q*’ to ‘necessarily, *p* if and only if *q*’ is an explanatory one; it is because *p* just is *q*, that necessarily *p* just is *q*. According to this view, then, generalized identity and modality are chained together in an explanatory way, and the explanatory gap is bridged because identity is necessary.

But how promising is this strategy? First, it relies on the hypothesis that for *each* and *every* essence claim, there is a generalized identity claim. So far, I have not been able to come up with a counterexample, but the view might be vulnerable from that particular angle. Second, one might think that this sophisticated version of (Finean) essentialism is at odds with the idea that essence actually *is* the source of necessity. *Prima facie*, essence does not appear to achieve a ‘modal-job’ in the process: they entail generalized identities which, in turn, provide the missing link. In response, essentialists can argue that the entailment from essence to necessity

---

<sup>76</sup> Essentialists might also say that it is only the grammatical surface of essentialist claims that leads to doubts regarding the connection between essence and necessity, in the sense that the generalized identity claim between the two was implied—and not made explicit.

via generalized identity is an explanatory one as well: essence explains/grounds identity conditions, and these latter explain/ground necessities. Ultimately, then, modality traces back to essence, because there is no (metaphysical) necessity without identity, and no identity without essence!

But regardless of a definitive answer to the matters just presented, there is a more pressing question: can Fine's opponents be satisfied with the claim that generalized identity bridges the gap between essence and modality? For the explanatory gap can be said to have moved from the level of essence to that of generalized identity. And so the skeptic can now ask: why does generalized identity give rise to metaphysical necessity? Recall that, with respect to essence, the explanatory gap is a challenge raised by skeptical philosophers who ask for explanations of *why* and *how* essence (understood as real definitions) can generate necessity. These philosophers argue that just because it is essential to some  $x$  that  $p$  it is not, by and in itself, sufficient to explain that it is therefore necessary that  $p$ : they ask for explanations regarding whether/why/how essence grounds necessity, and merely saying 'it's just what essences do!' will not dispel their doubts. Such philosophers do not see how necessity can find its source in something that is non-modal. If these difficulties are not taken seriously, then, ultimately, a modal posit (similar to the necessity of essences) will have to be in place to bridge the gap between generalized identity and necessity, and we will run in same the kind of difficulties I have presented in the precedent chapter.

In order to address this difficulty, we need to take a closer look at the mechanisms by which generalized identity operates to generate modality. To that end, we need

to do better than working with the presupposition that the link between the generalized identity and metaphysical necessity is less controversial than it is between essence and metaphysical necessity. So far, two different strategies have recently been pursued in the literature to address the explanatory gap challenge at the generalized identity level. The first one is discussed by Jessica Leech in her (2020), while the second one is due to Correia and Skiles in their (2019). I will now discuss the former in the following paragraphs, and devote the next section to the latter.

To the question of why generalized identity leads to necessity, Leech’s answer is, in substance, to say that generalized identity gives rise to necessity because it is itself necessary. Leech acknowledges that by saying that generalized identity is necessary, we have translated the problem faced by Finean essentialists about the necessity of essence: if—as we have seen—essentialists can hardly do any better than *presuppose* the necessity of essence (that is, to take that modal status as a mere ‘posit’), then why should we accept that generalized identity is necessary on different grounds? Leech’s answer is that, as opposed to essence, there is a *proof* of the necessity of generalized identity. Structurally speaking, the proof proceeds in the same way as Kripke’s proof of the necessity of objectual identity. With the version of *Leibniz’s Law* introduced in the previous section, we can derive the necessity of generalized identity as follows:

1.  $Fx \equiv_x Gx \supset (\phi \supset [Gx//Fx]\phi)$  [Leibniz’s Law for G.I.]
2.  $\Box(Fx \equiv_x Fx)$  [ $\Box$  of G. self-I.]
3.  $Fx \equiv_x Gx \supset (\Box(Fx \equiv_x Fx) \supset \Box(Fx \equiv_x Gx))$  [1, 2]

$$4. \quad Fx \equiv_x Gx \supset \Box(Fx \equiv_x Gx) \quad [2, 3]$$

We have proved that generalized identity is necessary.

Two questions arise. First, *how* is necessity ‘transferred’ from generalized identity to the biconditional? (*i.e.*, from ‘ $Fx \equiv_x Gx$ ’ to ‘ $\Box \forall x \Box(Fx \text{ iff } Gx)$ ’)? Answering this question amounts to explaining *how*, in detail, the gap between essence and modality is bridged. Second, *what* is the source of the necessity of generalized self-identity? Leech does not discuss the first question, but it can be answered as follows:

- |  |                                       |
|--|---------------------------------------|
| 1. $Fx \equiv_x Gx$                                    | [Assumption]                          |
| 2. $Fx \text{ iff } Fx$                                | [Tautology]                           |
| 3. $Fx \text{ iff } Gx$                                | [1, 2, Leibniz’s Law for G.I.]        |
| 4. $Fx \equiv_x Gx \supset Fx \text{ iff } Gx$         | [1, 3]                                |
| 5. $\models Fx \equiv_x Gx \supset Fx \text{ iff } Gx$ | [4]                                   |
| 6. $\Box(Fx \equiv_x Gx \supset Fx \text{ iff } Gx)$   | [5]                                   |
| 7. $Fx \equiv_x Gx \supset \Box(Fx \equiv_x Gx)$       | [Nec. of G.I.] <sup>77</sup>          |
| 8. $Fx \equiv_x Gx \supset \Box(Fx \text{ iff } Gx)$   | [6, 7, logical closure of ‘ $\Box$ ’] |

This proof substantially relies on both *Leibniz’s Law* and the logical closure of necessity. There are probably different ways to derive the necessity of the biconditional, but what matters here is that we have a (seemingly) better glimpse of

---

<sup>77</sup> See the proof on the previous page.

the underlying (logical) mechanisms by which generalized identity operates to bestow necessity.

The second question is, according to Leech, the most problematic one. For she argues that essentialists have traded the modal posit of essences (or of real definitions, for that matter) for the necessity of *generalized-self-identity*. Thus, just as Finean essentialists have difficulties accounting for the necessity of essences, Leech argues that the essentialist who appeals to generalized identity runs into troubles explaining—and therefore identifying the source of—the necessity of *generalized-self-identity*. The reason is that in order to uphold the sourcehood thesis (according to which essence is the source of *all* metaphysical necessity), there *must be* an essence-fact to explain the necessity of *generalized-self-identity*. What that essence-fact might be is a substantial matter, but assume, for the sake of the argument, that it is a fact about the essence of generalized identity:  $E(G.I.)(Fx \equiv_x Fx)$ . According to the framework considered here, this essentialist claim is translated as ' $(Fx \equiv_x Fx) \equiv_x (Fx \equiv_x Fx)$ '. As a consequence, that latter claim will be what is supposed to bridge the gap between ' $E(G.I.)(Fx \equiv_x Fx)$ ' and the necessity of  $(Fx \equiv_x Fx)$ . To illustrate, consider the following schema:

1.  $E(G.I.)(Fx \equiv_x Fx) \rightarrow (Fx \equiv_x Fx) \equiv_x (Fx \equiv_x Fx)$ ;
2.  $(Fx \equiv_x Fx) \equiv_x (Fx \equiv_x Fx) \rightarrow \Box(Fx \equiv_x Fx)$

Now, the situation turns out to be somewhat identical to that of the Finean essentialist's: instead of having to explain *why* ' $E(F)G \supset \Box(F)G$ ' is true, the



generalized identity essentialist now has to explain why  $(Fx \equiv_x Fx) \equiv_x (Fx \equiv_x Fx) \supset \Box(Fx \equiv_x Fx)$  is true. Our initial questions—namely, *how* and *why* does generalized identity give rise to necessity—remain *unanswered*. Leech explains that this is so because in order for essentialists to give an answer, they have to establish the necessity of  $(Fx \equiv_x Fx) \equiv_x (Fx \equiv_x Fx)$  via a proof in order to bridge the gap between  $(Fx \equiv_x Fx) \equiv_x (Fx \equiv_x Fx)$  and  $\Box(Fx \equiv_x Fx)$ ; otherwise the necessity of that generalized identity claim would be a *brute* modal posit. The problem is that in order for such a proof to succeed, it will itself substantially rely to the necessity of generalized self-identity. However, argues Leech, doing so would be viciously circular—*i.e.*, we would appeal to the necessity of generalized-self-identity in order to explain why generalized-self-identity is necessary. Leech concludes that essentialists find themselves entangled in a situation where there is no obvious way out.

One might think that this argument shows that essentialists *should not* appeal to generalized identity *at all* in order to bridge essence with necessity, because no significant progress has and can be made. That is, since the proof of the necessity of generalized identity crucially relies on a necessity that escapes the explanatory scope of the essentialist theory of modality, there is no reason to opt for that specific framework: one merely trades a problem for another one, and in either case the claim that *all* metaphysical necessities find their source in essence is eventually compromised.<sup>78</sup> So far, the only upshot is that there is a *proof* of the necessity of generalized identity—as opposed to having no proof of the necessity of essences—

---

<sup>78</sup> We can push this reasoning further by saying that even if we take the necessity of generalized identity as brute, we have not successfully overcome the necessity horn of Blackburn's dilemma: there remains that 'bad residual must'.

but this is certainly not enough to opt for the entailment from essence to generalized identity.

That being said, one might be tempted to argue that the necessity of generalized-self-identity does not find its source in essence, but rather in *logic*. To give an example, ' $\Box(Fx \equiv_x Fx)$ ' would simply find its source in ' $(Fx \equiv_x Fx)$ '—*i.e.*, the latter entailing the former in an explanatory way. Of course, the question of the gap resurfaces. Here, however, I am tempted to say that the skeptic's position would be rather difficult to sustain. For, logical truths and tautologies *just are* necessary: to be  $F$  just is to be  $F$ , and  $p$  just is  $p$ , period. It is necessarily so; how could it be otherwise? Yet even if there is presumably no gap at this level, the pressing difficulty is that by endorsing this view, essentialists would be committed to a framework that differs considerably from their own initial view for the following reasons.

First, essences no longer achieve the task that was expected of them, because necessity ultimately finds its ground in *logic*. That is, essence is the ground of identities, but these identities receive their modal status from logical truths. Thus, there is presumably *no need* to appeal to the entailment from essence to generalized identity, since essences are merely 'ornamental' as far as necessity is concerned. Second, the view is at odds with Fine's thesis according to which logical necessities form a subset of metaphysical necessities. Rather, the latter derives from the former, and metaphysical necessities are a subset of the logical ones. Finally, the view is a *logicist* one: all subject matters regarding necessity and identity are, in fact, purely logical in nature. Here, the worry is that this view does

not seem to be as substantial as Finean essentialism: necessity finds its source in trivial matters, and it moves from one level to another one thanks to Leibniz's Law. The difficulty here is that when we think about metaphysical necessity and its source, we seem to have a more 'worldly' conception of these phenomena.

At this point, I think that we can safely reject the logicist view, since endorsing it forces essentialists to throw the baby out with the bathwater. I will consider a different kind of logicist framework at the end of chapter 5. For now, let us turn to how Correia and Skiles intend to bridge the gap between generalized identity and modality, and how they can explain the necessity of generalized-identity statements.

#### 4.IV Correia and Skiles's essentialist framework

So far, we have focused solely on the explanatory gap problem with the defense of an entailment that goes from essence to generalized identity. As we have seen, opting for this strategy is not compelling since it leads to dead ends rather than genuine progress. In what follows, I will present Correia and Skiles's 'pure' generalized identity framework, and detail how their account aims to address the explanatory gap challenge as well as explaining the necessity of essences. The mere entailment from essence to generalized identity that I have just discussed does not support an entailment from generalized identity to the necessity of essence. That is, we do not have the following:

$$Fx \equiv_x Gx \rightarrow \Box E(F)G$$

In order to obtain that entailment, we need the stronger view advocated by Correia and Skiles (2019), namely that essence is to be understood (or analyzed) in terms of generalized identity. Since we already have ‘ $E(F)p \rightarrow Fx \equiv_x Gx$ ’, the next step is to defend the converse:

$$Fx \equiv_x Gx \rightarrow E(F)G$$

This amounts to saying that generalized identity is more *fundamental* than essence. Moreover, we need to distinguish between what Correia and Skiles label *full* vs. *partial* essence. Roughly, to say that it is essential to bachelorhood that whatever is a bachelor is an unmarried (adult) male arguably consists in spelling out the *full* essence of bachelorhood, because nothing more needs to be ‘added’ to define what being a bachelor is. In contrast, to say that it is essential to what it is to be Socrates that Socrates is human is (presumably) only a partial account of what being Socrates is. For there might be additional characteristics that are needed to fully account for what it is to be Socrates—*e.g.*, to originate from a specific zygote. For the scope of our discussion, I will only focus on the notion of full essence and use my examples accordingly. Accounting for this latter notion in terms of generalized identity is pretty straightforward (*ibid.*: 11):

**Full-Generic-Essence** Being  $F$  is what it is to be  $G$  in full iff: for a thing to be  $G$  is for it to be  $F$  (in symbols:  $Gx \equiv_x Fx$ ).

To be precise, the ‘what it is’ expression on the left-hand side of **Full-Generic-Essence** corresponds to the essentialist’s locution of essence. Thus, the reading of ‘what it is’ is equivalent to ‘being *F* is essential to being *G*’, ‘it’s essential to being *F* to be *G*’, or ‘it is true in virtue of the nature of what it is to be *F* to be *G*. So according to **Full-Generic-Essence**, being an unmarried adult human male is *what it is* to be a bachelor in full if, and only if: for a thing to be a bachelor just is for it to be an unmarried adult male. This latter claim fully determines what it takes for something to be a bachelor—*i.e.*, it provides both necessary and sufficient conditions for bachelorhood. Importantly, **Full-Generic-Essence** is itself a statement of generalized identity and, as Correia & Skiles explain, they use ‘iff:’ merely for the sake of readability (*ibid.*).

On Correia and Skiles’s view, essence is *defined* in terms of generalized identity. Thus, essence is no longer the fundamental notion that determines the identity conditions of entities and gives rise to necessity—as opposed to what essentialists (might) wish to defend. On this view, then, essence just is (definitionally speaking) generalized identity. This view is at odds with the Finean essentialist’s ‘spirit’, but the principal upshot is that we probably have a better understanding regarding how the notion essence behaves if its theory is spelled out in different terms. *Prima facie*, the notion of identity is less ‘mysterious’, if I may say, than Fine’s (allegedly) primitive notion of essence. That being said, we have just seen in the last section that the entailment strategy gives us little hope of explaining any necessity, since we would have to eventually appeal to the necessity of generalized-self-identity—which would remain unexplained. In response to this difficulty, Correia and Skiles offer an account of necessity for their ‘essence as generalized identity’ framework.

Their strategy relies on what they label the *extensional correlate* of a generalized identity. Basically, the idea is that the extensional correlates of ' $Fx \equiv_x Gx$ ' and ' $p \equiv q$ ' are ' $\forall x(Fx \text{ iff } Gx)$ ' and ' $p \text{ iff } q$ ', respectively. Let me briefly stress that these authors choose to tailor their response by treating generalized identities as propositions and not statements, but this will not substantially affect the following discussion and arguments.

Next, Correia and Skiles explain that every generalized identity and, consequently, every extensional correlate of a generalized identity, is metaphysically necessary. They formulate three different accounts of necessity, but I will mention only two of them (2019: 11):

**Strong account**      A proposition is metaphysically necessary *iff* it is a logical consequence of the true identities;

**Weak account**      A proposition is metaphysically necessary *iff* it is a logical consequence of the extensional correlates of the true identities.

According to the **Strong account**, generalized identities are necessary, because every true generalized identity is a logical consequence of itself—*i.e.*,  $p \equiv q \models (p \equiv q)$ , and so we have  $\Box(p \equiv q)$ . Next, we obtain  $\Box(p \text{ iff } q)$  with the following principle: for every generalized identity  $f$ , the extensional correlate of  $f$  is a logical consequence of  $f$ . Thus, the **Strong account** delivers necessities in two different ways: generalized identities establish their own necessity *directly*, and the necessity

of the extensional correlate is established *via* the additional principle just mentioned.

The **Weak account** provides the same results, but in a different way:  $\langle p \text{ iff } q \rangle$  is necessary since it is a logical consequence of itself. And  $\langle p \equiv q \rangle$  is also necessary, but with the help of two additional assumptions: (a) For all identities  $\langle \phi \equiv \psi \rangle$ , if  $\langle \phi \equiv \psi \rangle$  is true, then so is  $\langle (\alpha \equiv \phi) \equiv (\alpha \equiv \psi) \rangle$ , and (b)  $\langle \forall \alpha (\alpha \equiv \alpha) \rangle$  is valid (*ibid*: 12). Necessities are generated as follows:

- |   |                            |
|---|----------------------------|
| 1. $p \equiv q$                             | [Ass.]                     |
| 2. $(p \equiv p) \equiv (p \equiv q)$       | [1, (a), LL]               |
| 3. $(p \equiv p) \text{ iff } (p \equiv q)$ | [2, extensional correlate] |
| 4. $p \equiv p$                             | [(b), validity]            |
| 5. $p \equiv q$                             | [3, 4]                     |
| 6. $\Box(p \equiv q)$                       | [5]                        |

By contrast with the **Strong account**, the **Weak account** establishes the necessity of generalized identities—*i.e.*, essences—*indirectly*—via a proof and plausible assumptions.

Since both accounts provide the same results but differently, I think it is important to find out which one is the closest to the essentialist's framework. Interestingly, the **Weak account** is structurally similar to the explanations of necessity we have surveyed in chapter 3 regarding the necessity of essences as follows:

$$p \equiv q \rightarrow \Box(p \text{ iff } q)^{79}$$

$$E(x)p \rightarrow \Box p$$

$$(p \equiv p) \equiv (p \equiv q) \rightarrow \Box(p \equiv q)$$

$$E(x)E(x)p \rightarrow \Box E(x)p$$

And for the **Strong account**:

$$p \equiv q \rightarrow \Box(p \equiv q)$$

$$E(x)p \rightarrow \Box E(x)p$$

$$(p \equiv q) \rightarrow \Box(p \text{ iff } q)$$

$$E(x)p \rightarrow \Box p$$

According to the **Strong account**, essences make themselves necessary as well as their extensional correlates. Hence, the **Weak account** seems to be the candidate that corresponds the closest to how Finean essentialists can account for the necessity of essences—*i.e.*, with the appeal to iterated essences. So as opposed to the strategy defended in the previous section, it is apparently possible to account for the necessity of essences with the accounts propounded by Correia and Skiles.

But *how* is the explanatory gap problem addressed? Recall that skeptical philosophers ask for an explanation of how essence (understood here as generalized identity) gives rise to necessity. On either of the two accounts just presented, the answer of Correia and Skiles consists in saying that if a proposition,  $p$ , is a logical consequence of a specific set of proposition(s), then  $p$  is necessary. Before discussing this answer in detail, it does not seem that we can understand how essence (*i.e.*, generalized identity)—in and of itself—gives rise to necessity. This is so because Correia and Skiles's accounts appeal to the notion of logical

---

<sup>79</sup> To be precise,  $p \text{ iff } q$  is necessary because it follows from  $p \text{ iff } q$ , which is an extensional correlate of a generalized identity, namely:  $p$  just is  $q$ .



consequence to provide the desired results. In other words, it strikes me that statements of essence do nothing ‘on their own’. Rather, it looks *as if* logical consequence together with Leibniz’s Law are doing the explanatory heavy lifting. Of course, necessities are derived from a specific set of statements (the generalized identities), but that does not mean nor entail that these statements can provide the required modal ‘oomph’.

I think that Correia and Skiles’s reply fails, and is not sufficient to address the explanatory gap challenge. The first point I want to flag is that these authors turn the metaphysical questions of essence and modality into purely logical matters: substitution (for the **Weak account**), logical consequence, entailment, etc. But when we think of essence—as Fine and other philosophers before him have thought of it—it does not seem that we are after logical principles to explain why essence is the source of necessity, and how it generates the modal layers of reality. In fact, substitution *salva veritate* is a purely *intensional* phenomenon that destroys explanation (Kim 1993, Schaffer 2009): variation and covariation have no explanatory power, since they merely state that there are (necessary) patterns, but they do not explain *why* they hold (Maurin 2019). Relatedly, logical consequence is a notion that does not back a substantive kind of explanation. Logical consequence is indeed a key-notion behind entailments and valid arguments, but not a notion that supports—in and of itself—a *generative* kind of explanation like grounding. Of course, there is a *direction* at stake in the proofs considered so far, but when philosophers ask *what* the source of metaphysical necessity is, it does not seem that they are after a special category of proofs/arguments and derivations. The point is that there is a difference between saying that *p* is necessary because it is *grounded*

in essence, and saying that *p* is necessary because it follows logically from a set of proposition(s).

Second, since Correia and Skiles are willing to address the skeptic's challenge by offering an answer to it, I contend that they consider the gap between essence and necessity as a *substantial* difficulty. However, their answer to the gap is to *define* necessity, and doing so risks to deflating the notion of necessity altogether. Yet even if we set this problem (granted that it is a problem) aside and if we suppose that a definition of necessity is a reasonable view *per se*, doing so can presumably not be considered as a way to bridge the explanatory gap at stake. For there being a gap is what precisely undermines the possibility of a definition for a skeptic and, for that matter, for whoever thinks that there is an opaque link between two phenomena, as explained in chapter 2. Consider again the case of the mental and the physical: skeptics who think that there is a gap between these two notions will not take a definition of the mental in terms of the physical as an answer to their worries. Their primary worry is precisely that there are crucial aspects of the mental—namely, its subjective and phenomenological traits—that fail to be captured in purely physical terms; and *for* the skeptic that jeopardizes the possibility of a definition of the mental in terms of the physical. In the case of essentialism, skeptics argue that essence is not *sufficient* for necessity. So why should essence construed as generalized identity be any different?

Let us briefly turn to the notion of logical consequence. I said just before that logical consequence seemed to play an important role in Correia and Skiles's definitions—as if it is responsible for the modal 'oomph' in generating other

necessities—but that cannot be true. To see why, consider the **Strong account**: every essence is necessary because it follows from itself. And how is logical consequence supposed to do any specific work here? After all, if you think about it, *everything* follows from itself:  $p$  follows from  $p$ ,  $q$  follows from  $q$ , etc. However, self-entailment is not sufficient for necessity. So the only thing that matters in Correia & Skiles’s definitions is the notion of generalized identity, nothing else. In other words, generalized identity (must) *suffices* for necessity. As far as a providing an essentialist framework is concerned, this can work. But then these definitions fall short of addressing the explanatory gap challenge: *if*, as the skeptic thinks, essence *simpliciter* is not sufficient for necessity, then why should essence understood as generalized identity *simpliciter* be sufficient for necessity? Unless these authors have a specific kind of logical consequence in mind in their definitions, I do not see how their view is of any help to address the skeptic’s challenge. That is, we have no explanation as to why essence—understood as generalized identity—gives rise to necessity. Giving a definition strikes me as refusing to acknowledge the skeptic’s point regarding the gap between essence and necessity.

The resulting situation looks like a standoff. On the one hand, we have skeptics who are legitimately asking how and why essence/generalized identity gives rise to necessity. And on the other hand, there is Correia and Skiles who are willing to address the skeptic’s challenge, but they do so by offering a definition of necessity as an answer. It is as if these authors are saying that skeptics are ‘gapping’ the bridge between essence and necessity. All things being equal, I consider Correia and Skiles’s view as the generalized identity version of Correia’s (2006), where he

defines—in a Finean way—necessity in terms of essence. However, regardless of what the theoretical advantages of these views are, I do not consider them as sufficient to meet the skeptic’s challenge. By contrast, Leech outlines a much clearer view about the underlying mechanisms by which generalized identity operates to bestow necessity further. But as explained before, this comes at the price of a form of logicism together with brute necessities—something all essentialists alike should try to avoid.

Before concluding this chapter, we could, on behalf of Correia and Skiles, try to read the entailments in the **Strong account** and the **Weak account** of necessity as *non-definitional*. That is, as a substantial grounding structure:

**Weak account G:**  $p \equiv q < \Box(p \text{ iff } q)$ , and  $(p \equiv p) \equiv (p \equiv q) < \Box(p \equiv q)$ ;

**Strong account G:**  $p \equiv q < \Box(p \equiv q)$ , and  $\Box(p \equiv q) < \Box(p \text{ iff } q)$ .

Here, the structural similarity with non-reductive Finean essentialism is striking. Whatever option is available to the Finean essentialist seems to also be available to the generalized identity essentialist: defining necessity, grounding necessity, offering a theory of the source of necessity, and so on. However, the fact that generalized identity is so intimately close to the notion of essence actually plays against Correia and Skiles’s view. That is, the similarities with Finean essentialism come at the price of having to embrace the difficulties it faces as well. The two accounts above formulated in terms of grounding are no different—except for their syntactical form—than what a Finean essentialist would say:

$E(x)p < \Box p$ ,  $E(x)E(x)p < \Box E(x)p$ , and

$E(x)p < \Box E(x)p$ ,  $E(x)p < \Box p$

We saw in the previous chapter what the difficulties related to essences explaining their own necessity are, as well as the difficulties that follow from appealing to iterated essences. And abandoning Correia and Skiles's definitional stance towards modality just takes us back to square one with respect to these problems. Skeptics will not be convinced by these explanations, because their worries and questions are now formulated in terms of generalized identity: *how* and *why* does ' $p \equiv q$ ' give rise to ' $\Box(p \text{ iff } q)$ ', and so on.

To summarize this chapter, generalized identity is of no help for essentialists. Opting for the route where an entailment from essence to generalized identity is defended to bridge the gap leads us to difficulties with the necessity of generalized-self-identity. Either that necessity is to be left as brute on pain of circularity, or we need to opt for a form of logicism that takes us (way) too far from the core principles of Finean essentialism. Moreover, there is no straightforward way to account for the necessity of essences on this view. The other option, Correia and Skiles's essentialist framework, consists in providing *definitions* of necessity that rely on core logical principles. This is not satisfying because one does bridge an explanatory gap between two notions by defining one in terms of the other. Reading what these authors defend in a non-definitional way is also of no help since we merely find ourselves in the original essentialist's position, facing the

problems they face in the same way they encounter them. I therefore suggest that essentialists do not rely on generalized identity.

In the next chapter, I want to revisit Hale's non-reductive essentialism where the necessity of essence is fundamental, and see how some recent discussions in the literature might help us bridge the gap. Afterwards, I will give my final comments on essentialism and conclude the present investigation with my own take on the source of necessity.

## Chapter 5: Non-reductive essentialism

### revisited

Throughout the last three chapters, I have discussed two main difficulties for essentialists. The first one consists in an explanatory gap between essence and necessity. The second one pertains to a problem of extensionality, and I argued in chapter 3 that the necessity of essence cannot be successfully accounted for on pain of vicious regress and/or circular explanations. As we have seen, these difficulties cannot be overcome if essentialists opt for the Finean or Halean frameworks. Furthermore, it appears that a reductionist approach leaves no room for bridging the gap, since it basically amounts to saying that there never was a gap in the first place. Hale's non-reductivism also faces the explanatory gap challenge, and its proponents must work with a restriction on the scope of application of **Source**.

From there, we have tried to see whether a different framework—one where essence is not construed on the model of real definitions but rather analyzed in terms of generalized identity—could help us address these issues. Unfortunately, this approach does not only suffer from a series of problems that affect all reductionist theories of modality, but also falls short of addressing the main difficulties I have identified. In fact, it seems that, with the generalized identity framework, we have merely changed the words we use to name the things we are

working with, but we have not achieved any substantial advancement towards addressing the gap. As I have explained in the last chapter, the prospects of the generalized identity approach are dim, and I contend that any reductive interpretation of **Source** is doomed to fail with respect to the gap problem.

What I want to do now is to return to Hale's non-reductive essentialism and to simply bite the bullet regarding the necessity of essence. At this point, I think that it is preferable if essentialism can only explain a certain subset of the necessities—on the condition that there is a way to address the gap challenge. In this chapter then, I will follow Hale and assume that even if essences are necessary, this modal status is fundamental and it can therefore not be explained further. The structure of reality is, according to this view, one where there are (at least) two different layers of necessities: the fundamental essentialist ones from which the non-essentialist ones are derived. Remember that what is crucial for Hale's view is that essences provide a non-transmissive explanation of necessity, that is: even if  $E(x)p$  is necessary, it is its truth alone that explains  $p$ 's necessity. If non-transmissive explanations of necessity fail, the entire view collapses (at least under an essentialist perspective) because even though a necessity can explain another necessity, the question of the source will remain open. In other words, it is only if an essence's truth explains a necessity that we block the threat of a transmissive framework of necessity.<sup>80</sup> It is by keeping this in mind that I will now discuss the content of this chapter.

---

<sup>80</sup> Again, the matter here is that if  $\Box E(x)p$  explains  $\Box p$ , we have an explanation of  $\Box p$  but not of  $\Box E(x)p$ , since we cannot have explanations of the latter with the current assumption that the necessity of essences is fundamental. Thus, the transmissive framework will be nothing more than a view where some necessities are mapped with one another, but not an (essentialist) account of the source of necessity. I discuss this point in the final section of this chapter.



The plan is the following. In section 5.I, I characterize and explore in detail whether Hale's restricted view can bridge the gap based on a recent discussion by Wallner and Vaidya on this matter. Two main claims will be identified, and the remainder of this section will focus on the first of them. Section 5.II focuses on the second claim made by Wallner and Vaidya, and I will discuss a series of issues that arise when the necessity of essence is granted an explanatory power. Notably, I will argue that, as a matter of fact, there are several reasons to think that the necessity of essences plays a direct role in explaining derivative necessities. The general verdict from the two first sections will, again, be a negative one: Wallner and Vaidya's proposal cannot address the gap challenge, and is itself subject to a series of objections that directly target the idea that essences can explain necessity. In the third and last section, I will discuss some general ideas, and explain why I think that there are interesting alternatives to essentialism when questions of the source of necessity are seriously considered.

## 5.I Non-reductive essentialism vs. the gap: a reply to Wallner and Vaidya

Remember that the skeptic's question boils down to this: how can essence *simpliciter* be a sufficient ground for necessity? If we think again about Schaffer's H<sub>2</sub>O gap, grounding actually does the job because it operates within the *same* 'realm', within the *same kind* of facts about chemistry. The problem with essence and necessity is that the gap is a substantial one and that, in and of itself, does not

warrant a straightforward path towards reduction or (grounding) explanation. To that effect, the essence-necessity gap is similar to the gap in philosophy of mind. As we have seen in chapter 2, grounding does not help clarify the connection, because what the appeal to grounding does is allowing one to merely state that some facts about essence are correlated with some facts about necessities; that the former are more fundamental than the latter.<sup>81</sup> Grounding does not tell us *how* or *why* essence generates necessity. So we need to bring in something else or modify **Source** even further.

To that effect, I want to focus on Hale's views on essentialism and take the following characterization as our starting point:

**Non-reductive Haeian essentialism:** For every *non*-essentialist proposition  $p$ , if  $p$  is necessary, then  $p$  is necessary because for some  $x$ , it is essential to  $x$  that  $p$ .

This excludes all necessary propositions that contain an essentialist operator from falling within the explanatory scope of the theory. Next, I want to introduce the following claim:

**Fundamentality of the necessity of essence:** For every proposition  $p$ , if  $p$  is an essentialist claim of the form 'E(X)A', then  $p$  is a fundamental necessity (*i.e.*,  $p$ 's necessity cannot be explained).

---

<sup>81</sup> This is also seriously open to debate since we have merely assumed since the beginning that essence is the source of necessity. But from what I can tell, it could be anything else and essence might not even be more fundamental than necessity.

With these two claims, we completely shut down the necessity of essence's problem and we have a good grip on what Hale's view consists in. The explanatory gap is all we are left with. Importantly, we have to do better than what Hale defends in his works. Typically, the view Hale is advocating is something among the following lines:

1.  $\Box E(x)p$ ; (**Fundamentality of the necessity of essence**)
2.  $E(x)p$ ;
3.  $\Box p$ ;
4. (2) is what explains (3); (**Non-reductive Haeian essentialism**)
5. (1) plays no explanatory role regarding  $p$ 's modal status.

The main problem with this schema is that the explanatory gap holds between (2) and (3), and given what I have discussed in chapter 2, we have no straightforward way to bridge this gap. This is precisely what we have to address on pain of going back to square one. So, as has just been said, something needs to be added or modified in the above argument in order to address the gap problem.

To that effect, Wallner and Vaidya (2020) have recently argued that Hale's framework can be elaborated in such a way that it becomes possible to bridge the explanatory gap. They claim that modality is 'inherent' or 'built into' the notion of essence. Roughly, the idea is that we *cannot* have a non-modal source of necessity, and that applies to essence as well. So what does that mean, exactly, when these authors say that modality is 'built into' essence? Here is their explanation:

our preferred way of cashing out all these metaphors more precisely is to cast this in terms of essence: it is essential to Fine's notion of essence that essentialist truths hold necessarily.<sup>82</sup> (2020: 10)

Next, these authors explain that the fact that essences are necessary is sufficient for essences to have the *capacity* to explain necessity (*ibid.*: 12). In other words, it is because essences are necessary that they have the capacity to explain/ground necessity. What is crucial for Wallner and Vaidya is that  $\Box E(x)p$  does not enter into the explanation of  $\Box p$ . If that were to happen, they would have what Hale calls a *transmissive* explanation of necessity, and the idea that essence is the source of necessity would collapse, because the *explanans* of  $p$ 's necessity would itself be necessary. (Again, if non-transmissive explanations of necessity are possible, as they argue, then even though essences are necessary,  $\Box E(x)p$  should not be the *explanans* of  $p$ 's necessity on pain of compromising the distinctive explanatory role of  $E(x)p$  as *the* source of the necessity of its prejacent proposition.)

Aware of this important point, Wallner and Vaidya claim that the reason *why* essences can explain necessity differs from the *way* essences explain necessity (2020: 14-15). Their thoughts regarding this point can be schematized as follows:

---

<sup>82</sup> Let me briefly flag one issue here. These authors claim that it is essential to essence that essentialist truths are necessary and should be necessary if true. But how can this necessity be explained? Even though Wallner and Vaidya do not say explicitly whether the necessity of essence can be explained, it would play in their favor to take it as fundamental on pain of being entangled with the difficulties identified in chapter 3. Since I assume that the necessity of essence is fundamental here, I will not discuss this issue any further.

1. Essences are *essentially* necessary—i.e., it is essential to essence that if  $E(x)p$  then  $\Box E(x)p$ ;
2. The fact that essences are necessary explains why essences can explain necessity:  $[\Box E(x)p] < [[E(x)p] < [\Box p]]$ ;
3. However,  $[\Box E(x)p] < [[E(x)p] < [\Box p]]$  does not entail that  $[\Box E(x)p] < [\Box p]$ .

To be precise, the entire formula in (2) embodies Wallner and Vaidya's answer to the question of *why* essence can explain necessity, while the *way* essences do so is captured with ' $[E(x)p] < [\Box p]$ ' and clarified in (3). So the crux of Wallner and Vaidya's view is that  $\Box E(x)p$  and  $E(x)p$  play very distinctive explanatory roles. On the one hand, these authors explain in (2) that they can address the gap challenge because the necessity of essence (i.e.,  $\Box E(x)p$ ) is what explains the fact that  $E(x)p$  can explain the necessity of  $p$ . On the other hand, (3) is meant to clarify the idea that the necessity of essence does not play the role of the *explanans* regarding  $p$ 's necessity, since they endorse Hale's non-transmissive view of explanations of necessity, which requires that  $E(x)p$  is—and should be—the only relevant ground for  $p$ 's necessity.

So far so good. But if Wallner and Vaidya's explanations *did* succeed, they would consist in an important improvement towards the gap challenge. However, I think that this neo-Halean framework fails, and that saying that the necessity of essence has an explanatory power brings more difficulties than solutions for essentialism. So far, I have nothing to object to regarding (1), but I must say that (2) and (3) strike me as problematic. I will argue in the rest of this section that what these

authors claim raises a series of substantial difficulties and, most importantly, that it *does not* bridge the gap. In section 5.2, I will argue that, as opposed to what these authors think regarding (3),  $\Box E(x)p$  actually *enters* the explanation of  $p$ 's necessity, and that a view where  $\Box E(x)p$  explains  $\Box p$  is a more natural path to take than the one where  $E(x)p$  explains  $\Box p$ .

Let us start with (2). One of the main questions is why we should buy into the idea that essences have some sort of *capacity* to explain necessity because they are necessary. Put differently, if we want to know why essences *can* explain necessity, is the answer 'because essences are essentially necessary' a satisfying one? Perhaps it is. However, I am not sure how that is supposed to be compatible with the idea that essence is *the* source of necessity. My concern here is that it very much looks like the necessity of essences is explanatorily *relevant* in explanations of necessity, because that modal status *vindicates* the idea that essences *can* explain necessity—*i.e.*, an essence's modal status is therefore what confers to that essence the *ability* to explain necessity. But how is this not, in fact, *actually* saying that in order to explain necessity, an essence *relies* on a modal fact? It is difficult to pin down exactly what the problem is with requiring the necessity of essence to achieve an explanatory task, but to put it succinctly, it looks like it is a substantial *precondition* for Wallner and Vaidya's solution to go through.

Thus my general concern here is that we were initially told by essentialists that essence *simpliciter* is *the* ground of necessity; but now in light of recent criticisms about the gap, we are told by these authors that essence (*simpliciter*) *itself* relies on a modal fact in order to have the capacity to explain necessity. It sounds very much

like there is an underlying issue of *priority* in this novel essentialist structure of explanations—*i.e.*, the necessity of essence has to come first to secure the fact that essence can explain necessity, and I want now to identify some of the difficulties that follow from this idea.

Just like Hale, Wallner and Vaidya argue that *grounding* holds between a given essence and a necessity. So what about the explanatory link between the necessity of essence and the fact that a given essence explains a necessity? It seems to me that we are dealing with a *meta-grounding* question in this particular case—*i.e.*, a question of what grounds a given grounding claim. As a matter of fact, I think that it should be a grounding relation that holds between  $[\Box E(x)p]$  and  $[[E(x)p] < [\Box p]]$ , because the way I understand it, the necessity of essence *guarantees* and *makes it the case* that essence can explain necessity. This is a *substantial* claim and not some kind of ‘lightweight’ explanation. To illustrate, (2) can be translated as follows:

4.  $[\Box E(x)p] g < [[E(x)p] g < [\Box p]]$  (where ‘g<’ denotes grounding)

There are two main problems with (4). First, one might have doubts about whether facts about grounding are themselves grounded. After all, this is not a trivial matter and this question has occupied the mind of grounding theorists for over a decade now. Yet if there is indeed a grounding relation that holds between the necessity of essence and the grounding fact, there is a possibility that the necessity of essence is more fundamental than essence *simpliciter*. And I am not sure how that matches with a theory that wants to put essence first and make it the source of necessity

(even though the necessity of essence is taken to be fundamental). What I am saying here is that this scenario looks different from a view where *both* the necessity of essence and the essence *simpliciter* are fundamental, as I think it should be if we consider the pure Hlean view. If (4) is true and, for instance, one is sympathetic to Sider's *purity principle* (2013), it would mean that the constituents of ' $E(x)p$ ' are both fundamental and non-fundamental, because on the one hand, they constitute a fact that is fundamental—*i.e.*,  $[\Box E(x)p]$ —and on the other hand, they are part of a grounded (and therefore non-fundamental) fact—*i.e.*,  $[[E(x)p] g < [\Box p]]$ . This 'Schrödingerian' situation is something that I find odd.<sup>83</sup> Second and most importantly, even if grounding facts like  $[E(x)p < \Box p]$  are themselves grounded, there might be other—and perhaps 'better'—contenders than  $[\Box E(x)p]$  to provide this kind of explanation. For instance, one of the most popular views is Bennett's (2017), who argues for *Upwards Anti-Primitivism*—a view according to which the grounds ground that the grounds ground the grounded. In the present case, this would mean that  $[E(x)p]$  grounds that  $[E(x)p]$  grounds  $[\Box p]$ , which is in opposition to (4) above.

What I have just discussed above are merely some of the initial difficulties and questions that arise when an explanatory power is given to the necessity of essence. While I do not have a definitive answer to these concerns, it is clear that Wallner and Vaidya's view departs substantially from Hale's. For even though Hale says that the necessity of essence cannot be explained, he remains silent regarding

---

<sup>83</sup> There might be a way to actually maintain the idea that both  $[\Box E(x)p]$  and  $[E(x)p]$  are fundamental even though the latter explains a fact that involves the former.



whether the necessity of essence plays any kind of explanatory role.<sup>84</sup> However, when Wallner and Vaidya explain that the necessity of essence plays an *actual* explanatory role and that they are endorsing grounding for related matters, it raises alarm bells. Of course, they could clarify their view and argue that the explanatory relation at stake in (4) is not grounding, but this will look more and more like an unstable view that tries to seal every unexpected leak with what it has on hand.<sup>85</sup> So far then, my main concern is that granting an explanatory role to the necessity of essence is not a trivial matter, and as I see it, it is only considered in the present case to prove some kind of point against a substantial objection—the gap—and it comes with a series of difficulties that pertain to broader philosophical domains.

Next, the principal issue with (2) is: what does it actually achieve to say that the necessity of essence explains *why* essence *can* explain necessity? Can this help bridge the gap? We must be cautious here. The skeptic's main question specifically targets the explanatory connection between an essence and the necessity of its prejacent: *why* and *how* does essence explain necessity? As explained earlier, Wallner and Vaidya do indeed differentiate the reason *why* essences can explain necessities and the *way* they do so, but their defense of (2) *only* answers the *why*-question. So far then, we are told that essence has the capacity to explain necessity because it is itself necessary—by stipulation—and that is irrelevant to *how* essence achieves this task. So on the one hand, the necessity of essence explains how essence can explain necessity, which leads us to think that this pertains to explaining *p*'s necessity, but on the other hand, the necessity of essence cannot be

---

<sup>84</sup> In fact, since he believes that non-transmissive explanations of necessity *do* succeed, he is not even questioning the explanatory power of an essence's necessity or the gap.

<sup>85</sup> This is something I discussed in chapter 3 regarding the principle **NT**.

relevant to  $p$ 's necessity, because otherwise the non-transmissive element of the theory collapses, since only ' $E(x)p$ ' is supposed to achieve this. So, no matter how we wrap our minds around this theory, if we believe that ' $\Box E(x)p$ ' is explanatory irrelevant in bringing forth the necessity of  $p$ , then it does not help us understand *how* essence generates necessity.

In their answer to the gap problem, Wallner and Vaidya admit what I have just said. They ask: "*How* do essences explain necessity?", to which they reply "*they simply do!*" (2020: 15). I am certain that no one who appreciates fully what the gap challenge really is about would be convinced by this answer, which basically amounts to saying that skeptics do not understand what they are after when they raise the gap problem. Furthermore, when providing additional explanations to their answer to the gap, Wallner and Vaidya also claim that it "is built into the very notion of essence that essences are necessary *and* that essences explain necessity" (*ibid.*: 16). But just as I have explained throughout this investigation numerous times, one cannot simply stipulate that kind of (contestable) claim and then pretend that everything that follows from it is true, and that those who disagree with it have simply misunderstood them.

Let us think about what happens if we use the above strategy with the mind-body problem. If the skeptic asks how does the physical explain the mental and is answered 'oh, the physical simply does that! It is part of the nature of the physical to explain the mental, it's just the way it is!', physicalists have not solved anything at all, but merely dodged the problem. This is what, in my opinion, is 'tricky' with the (Finean) notion of essence, because by taking it as a primitive and not

endowing it with any criteria to identify what is essential to some  $x$ , we are free to do as we please. In other words, if you think that it is essential to physical events to explain mental events, you ‘just have to’ postulate the claim, and it turns out that you solve the mind-body problem. Similarly, if you want to dodge the gap problem for the theory of essence, you just have to say that it is essential to essence to explain necessity, and so on. But that still does not answer how essence explains necessity, unlike what Leech tried to achieve in what I discussed in chapter 4.

It is also clear that (3) does not quite achieve anything, since its only purpose is to make sure that Wallner and Vaidya do not fall into the trap of saying that the necessity of essence does not enter the explanation of  $p$ ’s necessity, which would otherwise compromise the Halean idea that essences provide non-transmissive kinds of explanation. Therefore, I rest my case regarding Wallner and Vaidya’s view: appealing to the necessity of essence to explain why essences can explain necessity is of no help when it comes to the gap and no progress has been made to that end. In fact, we are exactly in the same situation as Hale but with the necessity of essence that does some explaining. What Wallner and Vaidya do is basically agree that non-modal essences cannot explain necessity, but then they stumble on the same issue that haunts all other construals of essentialism, namely that they cannot provide any explanation regarding *how* essences *simpliciter* explain necessity: it is just a brute fact that they do it, and no skeptic would take that as a satisfying answer.

In the next section, I want to focus specifically on (3) and I will argue that, no matter whether my above criticisms and objections are sound or not, Wallner and

Vaidya's view is untenable, because if they grant an explanatory power to the necessity of essence—something that is totally new in the essentialist literature—there are reasons to think that a transmissive framework of explanations of necessity is more appealing than a non-transmissive one.

## 5.II On the necessity of essence's role in explanations of necessity

In (3), Wallner and Vaidya explain that just because the necessity of essence explains why essences can explain necessity, it does not follow that this modal status plays a role in explaining necessity. What these authors fail to appreciate, however, is that the necessity of essence clearly plays at the very least an *indirect* role in explaining necessity. Importantly, that modal status is needed not because essences are intuitively necessary, but because it has to achieve, in this context, a very specific task: explaining why essence can explain necessity. I could discuss the indirect role further, but I rather take a shortcut and go down deeper the rabbit hole by arguing that, in my opinion, the necessity of essence plays a *direct* role in explaining necessity on Wallner and Vaidya's view. The difficulty I want to raise is that as soon as we admit that the necessity of essence plays an explanatory role, it opens Pandora's box. In order to understand why, we need to focus on the following passages where they claim that essences are *essentially* necessary:

We are inclined to [say] that the fact that essences (or essentialist truths) are essentially necessary just means that essences – by their very essence – are modal posits. (2020: 11)

If the modal posit, i.e. the necessity, is essential to essence, then the essence is a modal posit and, hence, becomes a part of the modal family. (*Ibid.*)

The claim that essences are essentially necessary licenses talk of essences being modal, [...]. In light of the way we understand the notion of reduction in this context, this in turn shows why taking essences to be essentially necessary precludes a reductive analysis or reduction of modality in terms of essence. (*Ibid.*)

What these authors are saying is that essences cannot be ‘detached’ from their modal status—that their necessity is *built into* them, that they are part of the modal family by being essentially necessary. In the last quote, they explain that the fact that essences are essentially necessary is what prevents a *reduction* of modality to essence (for no reduction of necessity is possible if essence is itself necessary, something I entirely agree with).

However, here is the first issue with (3). If what they say were true, it should be a ‘two-way street’. The main question here is ‘why is (3) true?’. Put differently, how come is it that, in this particular case (as well as any instance of an essence explaining a necessity, for that matter), it is *just* the essence *simpliciter* that

provides the explanation of necessity—why is it that ‘ $E(x)p < \Box p$ ’? If we follow Wallner and Vaidya closely, it seems that we cannot dissociate ‘ $E(x)p$ ’ from ‘ $\Box E(x)p$ ’, since the latter is *essentially* built into the former. But if that is supposed to prevent reduction, then, I think that it also prevents ‘ $E(x)p$ ’ from being *the explanans* of ‘ $\Box p$ ’—which is in direct contradiction with what (3) is supposed to establish, namely that ‘ $\Box E(x)p$ ’ cannot be the *explanans* of the necessity of ‘ $p$ ’. So if essences are essentially necessary, how is it possible for the necessity of essence to not *directly* enter in the explanation of ‘ $\Box p$ ’? Think about it: on the one hand we are told that essence is so intimately connected to necessity that it is *essentially* necessary, but on the other hand, essence is somehow ‘detached’ from that modal status when it comes to explaining the necessity of ‘ $p$ ’. It strikes me as an arbitrary explanation whose sole motivation is to overcome the gap difficulty, and (2) together with (3) muddy the waters beyond what I take to be reasonably acceptable.

Let us elaborate further on the direct role that the necessity of essence can play on Wallner and Vaidya’s picture. Since they bestow an explanatory power to the necessity of essence to bridge the gap, I want to ask in what way—if any—is the following explanation not successful or even wrong:

$$5. [\Box E(x)p] < [\Box p]$$

This is the second problem that pertains to (3): we know that, via logic and proof, one necessity can explain another. Yet even though Wallner and Vaidya grant that  $[\Box E(x)p]$  has explanatory power, somehow this power is limited when it comes to explaining the necessity of  $p$ . It seems to me that (5) depicts a perfectly natural and

acceptable way to explain why  $p$  is necessary: if one already has the necessity of essence with an explanatory power in their framework *and* the derivative necessity, then why not consider something like (5)?

The problem, of course, is that (5) does not take on board any kind of consideration about what the source of necessity is. So the question here is, what is this ‘mysterious’ characteristic that prevents the necessity of essence from *directly* explaining the necessity of  $p$ ? That is, why, all of a sudden, do we need to ‘dissociate’ an essence from its (essential, inherent, built-into) modal status to explain the necessity of  $p$ ? It strikes me that (5) consists in a perfectly good explanation, but no one has considered this in print so far, since the primary focus was always narrowed down to questions about the source of necessity.

To illustrate this point further, imagine you briefly explain to a student that you need a kind of ‘map’ of the explanatory connections between a wide and diverse range of necessities (*i.e.*, the task consists in mapping a portion of the modal landscape, if you prefer). The student knows nothing about questions of the source of necessity and is not instructed about it, but knows enough to tell whether some facts are more fundamental than others. At some point, that student will likely encounter fundamental necessities like  $E(x)p$ ,  $E(y)q$ ,  $E(z)r$ , etc. as well as the following necessities  $p$ ,  $q$ , and  $r$ . Since there clearly *is* a common component—*i.e.*,  $p$ ,  $q$ , and  $r$ —between the fundamental necessities and the derivative ones, the student will likely be inclined to think that there is a one-to-one correspondence between  $\Box E(x)p$  and  $\Box p$ ,  $\Box E(y)q$  and  $\Box q$ , and  $\Box E(z)r$  and  $\Box r$ , and will presumably map an explanatory connection between them: they will say that  $p$ ,  $q$ ,

and  $r$  are necessary because  $E(x)p$ ,  $E(y)q$ ,  $E(z)r$  are necessary. However, if friends of Hale's view are right, that student is mistaken, and is not allowed to draw an explanatory symbol between, say,  $\Box E(x)p$  and  $\Box p$ . As the supervisor, you notice this mistake, and you have now to explain to the student why this is wrong. What do you say? If we side with essentialists, the answer will be: 'in addition to  $\Box E(x)p$ , there also is the essence fact *simpliciter* and this fact is *what* explains why  $p$  is necessary, not  $\Box E(x)p$ '. 'But why is that so?' asks the student; 'why is it not sufficient and acceptable to say that  $\Box E(x)p$  explains  $\Box p$ ?'.<sup>86</sup> If you are not in the business of defending that essence is the source of necessity, I find it hard to defend that  $E(x)p$  and not  $\Box E(x)p$  explains  $\Box p$ .

It seems to me that essentialists are in a pickle when giving satisfying answers to any of that student's questions. The kind of explanations that will follow will very likely invoke questions and matters that pertain to reduction, the source of necessity, Blackburn's dilemma, and so on. But none of these are satisfying answers if we take stock and compare what we achieve with either ' $[E(x)p] < [\Box p]$ ' or ' $[\Box E(x)p] < [\Box p]$ '. It turns out that all explanations given by the essentialists are meant to 'save their theory at any cost'. So the question for essentialists is why is the mere truth of an essentialist fact *stronger* than its own necessity when it comes to explaining the necessity of its prejacent proposition? Why do we need to ignore the necessity of essence to explain necessity?

---

<sup>86</sup> Another even more troublesome and legitimate question from the student would be: 'well, you told me before that  $E(x)p$  provides a Glazier-essentialist explanation of  $p$ , so it seems therefore natural to me to think that just like a truth explains another truth (e.g.,  $E(x)p$  explains  $p$ ), a modal truth explains another (e.g.,  $\Box E(x)p$  explains  $\Box p$ ).



This brings us to the third problem with (3). Since the necessity of essence has an explanatory power, essentialists can no longer say that essence *simpliciter* is a sufficient ground for necessity. For even if they stick to their guns and ignore the problem I just discussed, it remains true that, on their view, it is the fact that essence grounds necessity *together* with the fact that the necessity of essence grounds why essence can explain necessity that explains why essences explain necessity. And unless I fail to understand what Wallner and Vaidya say, it sounds very much like essence *simpliciter* is not a sufficient ground for necessity. The modal posit is needed, even though essence *simpliciter* supposedly explains necessity. All that Wallner and Vaidya's theory actually tells us is that we need the necessity of essence more than anything if we want to bridge the gap. Therefore, there clearly is an important difference between what essentialists defend, namely  $[E(x)p] < [\Box p]$ , and (5): the latter can be a sufficient ground for  $[\Box p]$ , but the former is apparently incapable of bringing forth  $[\Box p]$  on its own, if Wallner and Vaidya are right. This is a general concern for essentialism, because if the theory is supplemented with any additional criteria, it *ipso facto* entails that essence *simpliciter* is not a sufficient ground for explanations of necessity.

The fourth problem is that there is probably no definitive way, generally speaking, to conclusively establish that the necessity of essence is explanatory inert. To understand this point, we have to focus on what Glazier calls necessitarian explanations (2021). The idea is that, just as essence essentially explains its preadjacent proposition, it is argued that necessities explain by means of necessitating the propositions that fall within their scope. To illustrate, if it is necessary that  $p$ ,

then that explains  $p$ . So from the assumption that  $E(x)p$  is necessary and that essence grounds necessity we have:<sup>87</sup>

- |                                     |  |
|-------------------------------------|--|
| (6) $\Box E(x)p$                    | (Necessity of essence)                         |
| (7) $E(x)p < \Box p$                | (Instance of essence as the source of $\Box$ ) |
| (8) $\Box E(x)p \text{ } n < E(x)p$ | (Necessitarian explanation)                    |
| (9) $\Box E(x)p$ explains $\Box p$  | from 6, 7, 8                                   |

In this derivation, the necessity of essence does enter the explanation via transitivity. Of course, one might immediately argue that the kind of explanation at stake between (7) and (9) is different: one is grounding explanation, the other is necessitarian explanation.<sup>88</sup> But I am inclined to think that they both fall under the general ‘metaphysical explanation’ label, just as I have explained in chapter 3 when discussing essentialist explanations. In any case, since  $\Box E(x)p$  is explanatory relevant to  $E(x)p$ , and since  $E(x)p$  is itself explanatorily relevant to  $\Box p$ ,  $\Box E(x)p$  pertains in some way to  $p$ ’s necessity whether we like it or not—at least in the general and metaphysical great ‘scheme’ of things. I believe that it can hardly be denied that the explanatory connection is in the background, and that, somehow, the necessity of  $E(x)p$  plays a role. The main issue is that the necessity operator of the essence-truth is, in an explanatory sense, *prior* to the necessity of  $p$ , and this leaves the door open to skeptical considerations regarding the idea that essences *simpliciter* are sufficient to bring forth the necessity of their prejacents.

---

<sup>87</sup> For the sake of the argument, I ignore the explanatory gap challenge in the derivation below.

<sup>88</sup> Glazier (2021) makes it clear that necessitarian explanation is not a special case of grounding.

At this point, I think (Halean) essentialists need to address the elephant in the room. The pressing issue is that everything they say, all the explanations they give, and all the different kinds of justification they appeal to are only there to uphold the idea that non-transmissive explanations of necessity are possible and *must* succeed on pain of having the thesis that essence *simpliciter* is the source of necessity collapse (a claim that we have captured in the present chapter with **Non-Reductive Halean Essentialism**). Importantly, the claim that essence is the source of necessity has not been justified or argued for by essentialists; it is something I have taken for granted for the sake of this entire discussion. So as I see it, all these answers to bridge the gap, avoid the difficulties raised, and so on, are merely there to defend something that has not even been argued for in the very first place.

Essentialists are therefore in a difficult situation with Wallner and Vaidya's sophisticated version of Hale's non-reductive framework. In fact, the non-reductive characteristic of this view only brings more trouble to the table because as soon as it is admitted that essences are necessary and have an explanatory power, that modal status sticks around in all of their explanations. What essentialists could do is to conduct an in-depth investigation of the idea that essences are devoid of any kind of modal status at all (but that would be rather strange as stressed in the previous chapters). Independently from considerations about the explanatory gap challenge, there is the question of where, metaphysically speaking, should we place essences in the structure of reality in this case? Are they outside the modal realm? If so, then how do they 'cross' that boundary and give rise to necessities? I leave it up to essentialists to find the answers, but I am not expecting anything that could convince me to think, at this point, that **Source**—no matter how it is modified or

characterized—is a viable thesis. In the next section, I want to give some final words to distance myself from the idea that essence is the source of necessity, and bring up what I take to be a positive, efficient, and simpler account of how necessity can be generated.

### 5.III Essence and transmissive explanations of necessity

So far, it appears that the essentialist theory of modality is unable to deliver its promises and to overcome the difficulties it faces, even after biting the bullet and admitting that the necessity of essences cannot be explained. At this point, we should seriously consider whether essence is the best contender for a theory of the source of necessity. What feels especially wrong about essentialism is that its supporters think that essence being the source is somewhat unquestionable. Wallner and Vaidya tell us that “it’s just what essence does!” (2020), and there is, as far as I know it, no argument that essence—and not something else—is *the* notion responsible to explain the modal realm. Leech is right when she says that

once we reject an account of essence in terms of necessity and rethink our notion of essence anew, it becomes an open question whether we should expect essence to imply necessity [...]. The essentialist *should not be swayed by a hangover from a view they reject*” (2020: 5-6, emphasis mine).

Why did it not occur to Lewis, for instance, that essence was the prime candidate for the source of necessity? Why did the plurality of worlds come to his mind first? Why does Vetter (2015) advocate for potentialities as the source of all modalities and not essence?<sup>89</sup> We might not have answer to these questions. What matters, however, is that essentialism is not as promising as it advertised.

I want to open the following discussion with the situation where scientists discovered that water is, in fact, composed of H<sub>2</sub>O. This was accomplished thanks to strong empirical data and proofs; scientists did not merely say that water is H<sub>2</sub>O ‘out of nowhere’ without conclusive evidence or argumentation. Had they simply stipulated that water is H<sub>2</sub>O, and someone asked why water is not NaCl instead, they could have empirically proven why water is H<sub>2</sub>O and not NaCl. Of course, a philosopher’s job is substantially different, since empirical observation is (in most cases) indecisive when it comes to metaphysical theses like what the source of necessity is. Yet, even though philosophers are free to speculate on a myriad of things that pertain to reality and its structure (to name only these), they cannot afford to come up with such postulates without at least an argument to sustain them.

Essentialists claim that essence is the source of necessity without any premises, and this is why the skeptic’s gap objection is so direct to the said claim. To make another analogy with the sciences, scientists have postulated the existence of a

---

<sup>89</sup> Interestingly this is not a problem that seems to surface with the gap in philosophy of mind. For it is clear that there are psychophysical connections, and therefore the idea that the mental could originate from something else than the physical is, *prima facie*, less plausible than the idea that necessity could be generated from something that is not essence.

celestial object with a specific set of physical properties in order to explain a close-by phenomenon (think of Vulcan or Nemesis, for instance). This might solve part of a problem they are facing, and give a *temporary best explanation* scenario to decipher something they currently have answers to. But what about essentialists and their sourcehood claim? What kind of problem does the postulate ‘essence is the source of necessity’ actually solve in contemporary philosophy? Of course it answers the question ‘What is the source of necessity?’ but without more explanation and details, this claim is neither stronger nor more conclusive than ‘ $x$  is the source of necessity’. In fact, this theory cannot even explain the necessity of its own source, and when asked how and why essence explains necessity, we are left with, at best, an essentialist claim such as ‘it is essential to essence to explain necessity’.

In what follows, I want to explore the question of whether something different from essence could actually play the role of the source of necessity. After all, we are free to speculate at this point and my aim is to show that, with logic and a few plausible assumptions alone, we can have a theory of the source of necessity that uses essentialist truths, but in a way that has not been argued for before. I have flirted with the idea that a necessity can explain another necessity numerous times during this investigation, and I want to show now how this can be done, how a view where something like (5)—*i.e.*,  $\Box E(x)p < \Box p$ , can be defended.

To do so, I want to first consider the *distribution axiom* of modal logic:

$$\mathbf{DA}: \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$$

Basically, what this axiom tells us is that for every necessary conditional proposition, the necessity operator can be transferred to the constituents of the conditional. With this in mind, a principle like **NP** (*i.e.*, if  $F$  is an essential property of  $x$ , then  $F$  is a necessary property of  $x$ ) is also necessary if true at all. So what we have (reformulated here with propositions instead of properties) is the following principle:

$$\mathbf{NP}^*: E(x)p \rightarrow \Box p;$$

Now, **NP\*** is necessary if true, and with **DA** we can infer the following:

$$(1) \Box \mathbf{NP}^* \rightarrow (\Box E(x)p \rightarrow \Box p)$$

But it seems that the correct formalization is not (1), but rather:

$$(2) \Box \mathbf{NP}^* \rightarrow (\Box E(x)p \rightarrow \Box \Box p)$$

With (2), we still have the possibility that essence implies the necessity of its prejacent. But what I would like to ask now is whether essence is needed at all to play the distinctive role of the source of necessity. For consider the following: it is not merely true that essences imply their prejacent proposition, but necessarily true:

$$(3) \Box (E(x)p \rightarrow p) \quad (\text{Factivity of essence; every essence is a truth.})$$

With the distribution axiom, we can then infer that:

(4)  $\Box(E(x)p \rightarrow p) \rightarrow (\Box E(x)p \rightarrow \Box p)$       **DA**

(5) Therefore,  $(\Box E(x)p \rightarrow \Box p)$       (3, 4, *Modus Ponens*)

The ‘boxes’ in (5) are inherited from a necessary claim, (3), together with the application of modal logic’s distribution axiom (note that (3) is not an essentialist claim, but a necessary truth that *involves* a proposition prefixed with an essentialist operator). Moreover, note that essentialists should also be happy with the fact that there is no conflict with the idea that essences are necessary either, because essences *are* necessary on this picture (simply not how they intend it).

In the previous section, I briefly mentioned an alternative view—one where  $[\Box E(x)p] < [\Box p]$ —that I found very natural. Now, if what has just been said above with **DA** is true and if the derivation I just made holds, it seems that, on the picture where  $[\Box E(x)p] < [\Box p]$ , we might also need (3). That is, even if we simply assume something as simple as the necessity of essence to explain the necessity of their prejacent, it is still not a sufficient ground because the distribution axiom together with the factivity of essence principle are needed to *ensure* that necessity goes through. At the very least, we have now a story of *how* an essence and its prejacent proposition inherit their modal status.

Importantly, the arrows ( $\rightarrow$ ) in the above arguments do not denote explanations, but merely *implications*. So what this means is that there are two different layers at stake: the explanatory one and the implicative one. But this should not be a concern. As we know it, essentialist explanations hold between an essence and its



prejacent, and we also know that an essence being necessary can explain the necessity of its prejacent. We can identify a kind of general principle to that end:

**Truth and necessity:** if  $p < q$ , then  $\Box p < \Box q$

What the implication layer tells us, however, is *how* the necessity is transferred across an essence and its prejacent proposition. And that is done thanks to the distribution axiom. We already have the explanatory connections and the factivity of essence, but as it turns out, the necessity of essence alone is not sufficient to explain the necessity of its prejacent; the work of logic in the background with the distribution axiom is needed; otherwise we just have facts *simpliciter*. But thanks to logic the boxes are ‘attached’ to the essences and their prejacent propositions.

Ultimately then, necessity would find its source in something that is not essence—*i.e.*, logic—which differs from Hale’s and Fine’s views on the matter. On such views, essence is the alleged source from which necessity ‘flows’. But on this ‘Factivity-DA’ view I have just sketched, essence plays no distinctive role: it consists in one layer in metaphysical reality from which we can draw explanatory arrows towards the corresponding derivative necessities. This view is a ‘downstream’ picture of necessity that *attaches* itself to essence. And the source of all modality is ultimately flowing all the way down from **DA**. Anything that explains something else can transfer necessity across, granted that it is itself necessary, as I have just showed it with essence receiving its ‘box’ and transferring it further onto its prejacent. On the Finean essentialist view, essence is (supposed)

to *generate* necessity, while on the Factivity-**DA** view, essence plays the role of *transmitting* a necessity that it inherited.

I do not state this picture as an absolute truth, but rather as an eye-opening challenge for essentialists; there are *other*—perhaps even *better*—ways to give a theory of necessity without essence playing a substantial role in the process where it rather is a *part* of the said process. Of course, I expect the essentialist to resist all of the above claims. Perhaps they want to say that the source of the necessity at stake in the above picture is actually found in essence—perhaps something along the lines that the distribution axiom is essential to something. But, again, what can this something be? Essence *qua* essence? Does Essence know anything of the particular axioms of modal logic? And even if this were true, the explanatory gap would resurface at the level of the essence of essence: how can it be that essence bestows necessity upon an axiom? It seems to me that such axioms are necessary precisely because they are axioms, not because of something else. So perhaps it is the essence of the distribution axiom that is at stake, but this strikes me as utterly trivial and therefore uninformative:

(6) It is essential to the distribution axiom that the distribution axiom is true.

Fair enough, but if we follow the essentialist theory of necessity, it follows with (6) that we have the following:

(6\*) If E(**DA**)**DA**, then  $\Box(\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q))$

We still have no idea regarding how the first box in the consequent of (6\*) is generated. However, if we completely ignore the essentialist theory of necessity, we can start reasoning with one axiom of logic, and derive both the necessity of essence and the necessity of their preadjacent proposition. Such axioms are necessarily true, but the question of why they are necessary sounds confusing: if you understand what an axiom is, you understand that it is necessary. Axioms are necessary because they are axioms! Can a similar reasoning apply to essences? Of course, I can only expect essentialists to give an affirmative answer to this question: ‘If you understand what essential properties are, you understand that they are necessary!’. But at this point of the investigation it is worth stressing that this claim can be controversial: as Leech argues in her (2021), there are plausible ways to conceive of essences as not yielding necessity, which is hardly something that can truly be said of axioms.<sup>90</sup> The problem with explanations required for the necessity of essences seems to require from the essentialist’s theory that we appeal to another essence, and nothing else.

As a last remark, let me stress that it should be obvious at this point how **NP\***—*i.e.*,  $(E(x)p \rightarrow \Box p)$ —and (3)—*i.e.*,  $\Box(E(x)p \rightarrow p)$ —are importantly different. And I think that the latter is actually better than the former. The reason is that (3) does not involve anything controversial since the factivity of essence just is what it is: it states a truth. However, **NP\*** does not enjoy that status because as we have seen, it is actually controversial that essence should imply necessity (*i.e.*, the gap). Since

---

<sup>90</sup> Leech argues that, “insofar as we understand what the notion of essence is supposed to offer us, it can do that without having to yield necessity: necessity is not required” (2021: 17). To that effect, the view I have sketched in this chapter is stronger than the logicist view I presented in chapter 4.

(3) is more neutral with respect to its component claim, and since the necessity of the said components can be derived with minimal non-essentialist assumptions and principles, I would rather opt for a clear-cut framework like this one over something akin to **NP\*** to theorize about the source of necessity. I think that one of the upshots that essentialists might appreciate is that essence is still involved—and perhaps even needed—in this metaphysical process.

What this discussion so far should teach us is that essentialists really need to provide us good reasons to think that essence is *the* source of necessity, especially when such a simple framework like the one I have just sketched seems to work quite naturally and efficiently. It does not appear that essence is doing anything special on its own to qualify as the source of necessity, or to distinguish it from other possible kinds of explanations (except for what the essentialist wants us to believe). Importantly, what we tend to overlook is that if the essentialist theory of necessity is true, it should, I believe, rule out all other possible candidate explanations. It should be difficult to come up with another competitive framework or even to achieve any kind of explanatory results that could match what the essentialist is trying to do. However, it appears that essentialism is far more troublesome to defend than what its proponents think. Modality is, I believe, a rather strange and ‘elusive’ phenomenon that does not seem to owe anything to anything—it is just there for us to unveil—but crossing its metaphysical boundaries (should there be any) is something that hardly makes sense; at least for now.

# Bibliography

- Adams, R. M. (1981). Actualism and thisness. *Synthese*, 49, 3-41.
- Blackburn, S. (1986). Morals and modals. In *Essays in quasi-realism* (pp. 52–74).  
Oxford: Oxford University Press.
- Block, N., & Stalnaker, R. (1999). Conceptual analysis, dualism, and the  
explanatory gap. *The Philosophical Review*, 108(1), 1-46.
- Bovey, G. (2021). Essence, modality, and intrinsicity. *Synthese*, 198(8),  
7715-7737.
- Bovey, G. (2022). On the necessity of essence. *Philosophical Studies*, 179(7),  
2167-2185.
- Cameron, R. (2010). On the source of necessity. In B. Hale & A. Hoffman (Eds.),  
*Modality, Metaphysics, logic, and epistemology* (pp. 137–151). Oxford: OUP.
- Casullo, A. (2020). Essence and explanation. *Metaphysics*, 2, 88–96.
- Chalmers, D. J., & Jackson, F. (2001). Conceptual analysis and reductive  
explanation. *The Philosophical Review*, 110(3), 315-360.
- Chalmers, D. (2007). The hard problem of consciousness. In M. Velmans, & S.  
Schneider (Eds.), *The blackwell companion to consciousness* (pp. 225–  
235). Blackwell.
- Cohen, S. M. (1978). Essentialism in Aristotle. *The Review of Metaphysics*:  
387-405.
- Correia, F. (2006). Generic essence, objectual essence, and modality. *Noûs*,  
753-767.

- Correia, F. (2007). (Finean) Essence and (Priorean) Modality. *Dialectica*, 61(1), 63-84.
- Correia, F. (2012). On the reduction of necessity to essence. *Philosophy and Phenomenological Research*, 84(3), 639-653.
- Correia, F. (2017). Real definitions. *Philosophical Issues*, 27(1), 52-73.
- Correia, F., & Skiles, A. (2019). Grounding, essence, and identity. *Philosophy and Phenomenological Research*, 98(3), 642–670.
- Correia, F., & Skiles, A. (2022). Essence, modality, and identity. *Mind*, 131(524), 1279-1302.
- Cowling, S. (2013). The modal view of essence. *Canadian Journal of Philosophy*, 43(2), 248-266.
- Dasgupta, S. (2014). The possibility of physicalism. *The Journal of Philosophy*, 111(9/10), 557–592.
- Dasgupta, S. (2016). Metaphysical rationalism. *Noûs*, 50(2), 379-418.
- Denby, D. (2014). Essence and Intrinsicity. In R. Francescotti (Ed.), *Companion to Intrinsic Properties* (pp. 87-109). Berlin: De Gruyter.
- Dorr, C. (2016). To Be F is To Be G. *Philosophical Perspectives*, 30, 39-134.
- Dummett, M. (1959). Wittgenstein's philosophy of mathematics. *The Philosophical Review*, 68(3), 324-348.
- Fine, K. (1994). Essence and Modality. *Philosophical perspectives*, 8, 1-16.
- Fine, K. (1995). The logic of essence. *Journal of Philosophical Logic*, 24(3), 241-273.
- Fine, K. (2015). Unified foundations for essence and ground. *Journal of the American Philosophical Association*, 1(2), 296-311.
- Fitch, G. W. (2004). *Saul Kripke*. Chesham: Acumen.

- Glazier, M. (2017). Essentialist explanation. *Philosophical Studies*, 174(11), 2871–2889.
- Glazier, M. (2021). The difference between epistemic and metaphysical necessity. *Synthese*, 198(6), 1409-1424.
- Hale, B. (2013). *Necessary beings. An essay on ontology, modality, and the relations between them*. Oxford: OUP.
- Hale, B. (2018). The basis of necessity and possibility. *Royal Institute of Philosophy Supplements*, 82, 109–138.
- Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, 32(127), 127-136.
- Kaplan, D. (1986). Opacity. In S. Hahn & P. Schilpp (Eds.), *The Philosophy of W. V. Quine* (pp. 229-289). La Salle: Open Court.
- Kim, J. (1993). *Supervenience and mind: Selected philosophical essays*. Cambridge University Press.
- Kment, B. 2017. Varieties of Modality. In *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (Ed.), <https://plato.stanford.edu/archives/spr2017/entries/modality-varieties>.
- Koltzenburg, M., & Handwerker, H. O. (1994). Differential ability of human cutaneous nociceptors to signal mechanical pain and to produce vasodilatation. *Journal of Neuroscience*, 14(3), 1756-1765.
- Kripke, S. (1972). *Naming and Necessity*. Cambridge: Harvard University Press.
- Leech, J. (2018). Essence and mere necessity. *Royal Institute of Philosophy Supplements*, 82, 309-332.
- Leech, J. (2020). From essence to necessity via identity. Forthcoming in *Mind*. <https://doi.org/10.1093/mind/fzaa012>

- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific philosophical quarterly*, 64(4), 354-361.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell Publishing.
- Lowe, E. J. (2008). Two notions of being: Entity and essence. *Royal Institute of Philosophy Supplements*, 62, 23–48.
- Linnebo, Ø. (2014). ‘Just is’-statements as generalized identities. *Inquiry*, 57(4), 466-482.
- Lowe, E. J. (2012). Essence and ontology. In L. Novak, D. D. Novotny, P. Sousedik, & D. Svoboda (Eds.), *Metaphysics: Aristotelian, scholastic, analytic* (pp. 93–111). De Gruyter.
- Mackie, P. (2006). *How Things Might Have Been: Individuals, Kinds, and Essential Properties*. Oxford: OUP.
- Mackie, P. (2020). Can metaphysical modality be based in essence? In M. Dumitru (Ed.), *Metaphysics, meaning, and modality: themes from Kit Fine* (pp. 247–264). Oxford: OUP.
- Marcus, R. B. (1971). Essential Attribution. *The Journal of Philosophy*, 68(7), 187-202.
- Matthews, G. B. (1990). Aristotelian Essentialism. *Philosophy and Phenomenological Research*, 50, 251-262.
- Maurin, A. S. (2019). Grounding and metaphysical explanation: It’s complicated. *Philosophical Studies*, 176(6), 1573–1594.
- Moore, G. E. 2004. *Principia Ethica*. Mineola, NY: Dover Publications.
- Nagel, T. (1974). What is it like to be a bat?. *The philosophical review*, 83(4), 435-450.



- Neale, S. (2000). On a Milestone of Empiricism. In A. Orenstein & P. Kotatko (Eds.), *Knowledge, Language and Logic: Questions for Quine* (pp. 237-346). Dordrecht: Springer.
- Noonan, H. (2019). The new Aristotelian essentialists. *Metaphysica*, 19, 87–93.
- Papineau, D. (2011). What exactly is the explanatory gap?. *Philosophia*, 39, 5-19.
- Plato (1921). *Theatetus* (Translated by H. N. Flower). New York: William Heinemann.
- Plantinga, A. (1978). *The Nature of Necessity*. Oxford: OUP.
- Plantinga, A. (1983). On existentialism. *Philosophical Studies*, 44(1), 1-20.
- Quine, W. V. O. (1960). *Word and object*. Cambridge: MIT press.
- Quine, W. V. O. (1961). Reference and Modality. In W. V. O. Quine (Ed.), *From a Logical Point of View* (pp. 139-159). New York: Harper and Row.
- Rabin, G. O., & Rabern, B. (2016). Well founding grounding grounding. *Journal of Philosophical Logic*, 45, 349–379.
- Rayo, A. (2013). *The construction of logical space*. OUP Oxford.
- Roca-Royes, S. (2011). Essential Properties and Individual Essences. *Philosophy Compass*, 6(1), 65-77.
- Romero, C. (2019). Modality is not explainable by essence. *Philosophical Quarterly*, 69, 121–141.
- Rosen, G. (2010). Metaphysical dependence: Grounding and reduction, In B. Hale & A. Hoffman (Eds.), *Modality: Metaphysics, logic, and epistemology* (pp. 109–136). Oxford: Oxford University Press.
- Russell, B. (2020). *The principles of mathematics*. Routledge.
- Salmon, N. (1981). *Reference and Essence*. Princeton: Princeton University Press.
- Sayre-McCord, Geoff. 2011. Moral Realism. In *The Stanford Encyclopedia of*

- Philosophy*, Edward N. Zalta (Ed.). <http://plato.stanford.edu/archives/sum2011/entries/moral-realism/>.
- Schaffer, J. (2009). On What Grounds What. In *Metametaphysics*, D. Chalmers, D. Manley, and R. Wasserman (Eds.), pp. 347–83. Oxford: OUP.
- Schaffer, J. (Forth.). The ground between the gaps. *Philosopher's Imprint*.
- Sider, T. (2013). *Writing the Book of the World*. Oxford: OUP.
- Soames, S. (2004). *Reference and Description: The Case Against Two-Dimensionalism*. Princeton: Princeton University Press.
- Soames, S. (2005). Naming and Asserting. In Z. G. Szabó (Ed.), *Semantics versus Pragmatics* (pp. 356–382). Oxford: OUP.
- Stalnaker, R. (1979). Anti-Essentialism. *Midwest Studies in Philosophy*, 4(1), 343–355.
- Van Cleve, J. (1999). *Problems from Kant*. Oxford: OUP.
- Van Cleve, J. (2018). Brute necessity. *Philosophy Compass*, 13(9), 1–43.
- Vetter, B. (2015). *Potentiality: From dispositions to modality*. Oxford: OUP.
- Wallner, M. (2019). The structure of essentialist explanations of necessity. *Thought: A Journal of Philosophy*, 9(1), 4–13.
- Wallner, M., & Vaidya, A. (2020). Essence, explanation, and modality. *Philosophy*, 95(4), 419–445.
- White, N. P. (1972). Origins of Aristotle's Essentialism. *The Review of Metaphysics*, 26(1), 57–85.
- Wildman, N. (2013). Modality, Sparsity, and Essence. *The Philosophical Quarterly*, 63(253), 760–782.

- Wildman, N. (2018). On shaky ground? Exploring the contingent fundamentality thesis. In R. Bliss & G. Priest (*Eds.*), *Reality and its structure: Essays on fundamentality* (pp. 275–290). Oxford: OUP.
- Wilsch, T. (2017). Sophisticated Modal Primitivism. *Philosophical Issues*, 27(1), 428-448.
- Yablo, S. (1987). Identity, Essence, and Indiscernibility. *The Journal of Philosophy*, 84(6), 293-314.
- Zalta, E. N. (2006). Essence and Modality. *Mind*, 115(459), 659-694.
- Zylstra, J. (2019). Making semantics for essence. *Inquiry*, 62(8), 859–876.