How Theoretical Terms Effectively Refer

Sébastien Rivat*

February 9, 2025

Abstract

Scientific realists with traditional semantic inclinations are often pressed to explain away the distinguished series of referential failures that seem to plague our best past science. As recent debates make it particularly vivid, a central challenge is to find a reliable and principled way to assess referential success at the time a theory is still a live concern. In this paper, I argue that this is best done in the case of physics by examining whether the putative referent of a term is specifiable within the limited domain delineated by the range of parameters over which the theory at stake is empirically accurate. I first implement this selective principle into a general account of reference, building on Stathis Psillos's works. Then, I show that this account offers a remarkably reliable basis to assess referential success before theory change in the case of effective theories. Finally, I briefly show that this account still works well with other physical examples and explain how it helps us to handle problematic cases in the history of physical sciences.

1 Introduction

Many scientific realists share the intuition that our best current theories are approximately true. Their remarkable predictive success, it is usually said, would be hard to explain if they had nothing to do with what the world is like beyond phenomena. Yet this long-standing intuition has been challenged on many counts and perhaps most strikingly by the apparent twists and turns of the history of science. Seen by the light of their successors, some of our best past theories appear to be radically false and to contain central theoretical terms that fail to refer to anything real (e.g., 'gravitational force', 'phlogiston', 'luminiferous ether'). This suggests, first, that the explanatory

---

*Munich Center for Mathematical Philosophy, LMU Munich. Email: sebastien.rivat@lmu.de

link between predictive success, approximate truth, and reference dear to scientific realists might not be as robust as originally thought, and, second, that our best current theories might prove one day to be as radically false and referentially unsuccessful as some of their predecessors.

The most popular response in the realist camp is to concede that our best past theories did not get everything right but still maintain that some of their central parts survived and thus remain worthy of realist commitments (e.g., Kitcher, 1993; Psillos, 1999; Chakravartty, 2007). The "selective realists" who stick most closely to the traditional realist position, such as Kitcher and Psillos, also usually acknowledge that the specific problem of referential failure across theory change forces us to adjust both our semantic and epistemic commitments. In particular, if we grant that at least some terms in our best past theories fail to refer, we cannot just take the terms of successful theories to automatically pick out the right sorts of entities (as in causal or causal-historical theories of reference) and restrict ourselves to selecting descriptions that we can trust. We also need to account for: (i) the mechanism by which some, but not all, terms come to refer to unobservable entities; and (ii) the putative referential stability of some, but not all, terms under theory change (or their putative referential continuity if the domains of successive theories overlap).[1]

The central challenge underlying both (i) and (ii) is to find a reliable and principled way of distinguishing between referential success and failure, and this is far from trivial. For instance, it does not seem that we can rely on the theoretical content of our best current theories to assess past referential success since we do not yet know whether they will not appear to be deeply mistaken by the light of future theories. We also seem to be in the same situation with respect to current scientists' judgments insofar as their descendants might prove them wrong. And it does not seem that we can point to the crucial predictive and explanatory role of a term either since the next theory might show that, ultimately, this term was not playing such a crucial role. The challenge, in other words, is to find a reliable principle of selective reference, which works well even at the time the theory at stake is still a live concern, and adjust the semantics of theoretical terms accordingly.

---

[1] Structural realists would probably point out here that the amount of discontinuity is much less important once we focus on the (non-redundant) structural content of our best past theories (e.g., Worrall, 2020; Ladyman, 2021). Even if this proves to be right, both epistemic and ontic structural realists would probably still benefit from providing a robust account of: (i) the mechanism by which some, but not all, mathematical equations or structures come to relate to their target; and (ii) why some, but not all, mathematical equations or structures are likely to be referentially or representationally stable under theory change. For simplicity, I will restrict myself to selective strategies closely associated with the traditional realist position in this paper. As a side note, I will also use a broad notion of theory in the sequel (e.g., a single hypothesis, a theoretical model, a comprehensive theoretical framework together with a set of models).

I will refer to this challenge as "Stanford's challenge" following Stanford's (2006) criticism of the selective realist strategy with respect to the problem of referential failure (see also, e.g., Stanford, 2003a,b; Saatsi et al., 2009; Stanford, 2015).

My goal in this paper is to show that recent developments in theoretical physics offer a way to make some progress toward this challenge without making referential success too easy or too hard to achieve. I will first engage with Psillos's (1999; 2012) account of reference, which still constitutes, in my view, one of the best extant attempts to combine the lessons of the causal and descriptive theories of reference. Despite all its merits, I will argue that his account is nonetheless not fully satisfactory—although not exactly for the reasons usually raised in the literature (e.g., Stanford, 2006; Chakravartty, 2007).[2] In particular, I will argue that his specific use of causal (or causal-historical) and descriptive elements of reference fixing does not allow us to restrict appropriately the set of putative referents of a term across scales. (I should emphasize here that I will be working with a rather ontologically permissive and traditional metaphysical picture, in which the world does contain higher-level items like atoms and nuclei besides fundamental ones, if there is indeed any fundamental level.)

Then, building on Psillos's work, I will propose a new account of reference inspired by the framework of effective theories to address this issue. The key idea is to select terms according to whether their putative referent is specifiable within the limited domain of unobservable entities delineated by the range of parameters over which the theory at stake is empirically accurate (or, for short, specifiable within the "empirical reach" of the theory). I will modify Psillos's account accordingly, focusing on singular terms and treating general terms as such for simplicity, and show that the resulting account provides a more adequate basis for addressing Stanford's challenge in the case of physics. As we will see with toy models of Galilean and Newtonian gravitation, the selection process is remarkably reliable in the case of effective theories. I will also briefly explain how the account works with other physical examples and helps us to handle the usual suspects, such as 'luminiferous ether' and 'phlogiston' (a more liberal notion of range of parameters is required in the latter case). And, overall, the lesson of effective theories will be twofold: (i) typical cases of referential failure arise when the putative referent of a term in a theory lies beyond its empirical reach; (ii) there are good reasons to believe that the reference of a term is stable if the core causal-explanatory description of the referent does not depend significantly on a large variety of plausible and more comprehensive alternatives.

---

[2] For other proposals and responses that I will not discuss for lack of space, see, e.g., Saatsi (2005); Votsis (2011); Hoefer and Martí (2020); Vickers (2022).

There are several motivations for taking effective theories as a guide. First, our best current physical theories, i.e., the theories for which the problem of referential failure is arguably the most pressing, are widely believed to be most reliably formulated as effective theories (e.g., Weinberg, 1999; Levi, 2023). Second, the framework of effective theories comes along with powerful resources to assess whether we can speak reliably about remote or elusive entities (as we will see in section 4). Third, the concepts and methods of effective theories, including renormalization group (RG) methods, have been successfully extended to most areas of contemporary physics and found many applications beyond (including in quantum chemistry and molecular biology). And so the hope is that we might be able to learn something valuable from effective theories about the topic of reference beyond fundamental physics, and even physics itself (although I will have only very little to say about this here).[3] This paper should be read in that respect as part of the series of recent attempts to explore the implications of effective theories for the debate over scientific realism (e.g., Fraser, 2018; Williams, 2019; Rivat, 2021; Koberinski and Fraser, 2023; Dougherty, forthcoming). However, as I have already engaged with the more specialized part of this debate, I will not delve again into the details of effective field theories (EFTs) and RG methods here.

The paper is organized as follows. Section 2 discusses Psillos's works. Section 3 outlines the account of reference. Section 4 shows how it works in the case of effective theories. Section 5 extends the discussion to other examples and problematic historical cases in the physical sciences.

2 Psillos on reference

The debate over the reference of scientific theoretical terms is usually driven by two competing intuitions about reference fixing.

According to descriptivists, a term refers if and only if there is some unique entity satisfying the core description associated with the term (see Frege, 1892/1980; Russell, 1905, for early accounts). This, however, raises well-known issues for selective realists. For instance, if we require entities to be picked out by means of a comprehensive set of properties, we are likely to find incompatible sets over time and make the history of science more referentially discontinuous than it appears to us. Inversely, if we keep only a minimal set of properties to avoid this issue, we are likely to find that the core description associated with each term is satisfied by entities with otherwise

---

[3] To be clear: I am not suggesting to formulate every past and present scientific theory as an effective theory. Effective theories should only be considered as paradigmatic examples for which the general theory of reference presented in the paper works remarkably well.

radically incompatible properties, which would leave us with serious ambiguities as to what makes our most successful theories approximately true. And, as we will see with Psillos's descriptivist constraint below, descriptive theories of reference typically fail to provide a reliable and principled way of identifying the right amount of properties.

Causal (or causal-historical) theorists, by contrast, take reference to be originally fixed by means of some kind of causal contact with the entity for which a new term is introduced (or an old term reused). Competent speakers might attribute incompatible properties to this entity over time. But this does not introduce any referential discontinuity if they use the term in the same way and intend to refer to the entity originally picked out (see Kripke, 1972; Putnam, 1975, for early accounts).[4] As is well known again, this account makes referential success too easy to achieve: a term introduced to pick out some causal origin for a given observed phenomenon automatically refers to something insofar as there is presumably always some causal agent for any set of phenomena. This, however, leaves selective realists tempted by the causal theory in an unstable position. For it means that referential success does not depend at all on whether the theory at stake is approximately true and that we may, for instance, successfully speak about fundamental physical entities even if we are completely wrong about their behavior and core properties.

Psillos (1999; 2012) proposes to address these issues by taking the best of both worlds, and his account of reference can be reconstructed as follows. A term t in a theory T refers to an entity x under three conditions:

(C) Causal link: t is introduced (or used) to pick out some causal origin x for a set of observed phenomena ϕ;

(S) Satisfaction link: x satisfies the core causal-explanatory description of ϕ associated with t in the theory T;

(T) Tracking: The core causal-explanatory description of ϕ captures the set of kind-constitutive properties of x that play an indispensable role with respect to T in the causal explanation of ϕ.

By 'core causal-explanatory description' Psillos means the description of x that anything has to satisfy in order to play the same causal role as x with

---

[4] Strictly speaking, it is not part of Kripke's and Putnam's views that the reference-fixing event necessarily involves some causal contact with the referent of a term. For simplicity, I restrict myself to the "full causal theory of reference", to use Kroon's expression (1985, p. 144), which involves causal contact both in the reference-fixing event (e.g., by pointing at some target) and the reference-borrowing process (e.g., by learning how to use a term from another speaker). I also assume that indexical existential descriptive statements of the form 'there is an x that caused *this* phenomenon' involve causal contact with the observed phenomenon and thus indirectly with whatever is responsible for it.

respect to ϕ, and, by 'indispensable', that the kind-constitutive properties cannot be replaced by other non-ad hoc properties in T playing the same role in the causal explanation of ϕ (cf. Psillos, 1999, pp. 109-10, 294-5).

The most appealing aspect of Psillos's account, in my sense, is its precise identification of the shared burden of reference fixing via causal contact and via satisfaction of a description. We can always obtain some referent by stipulating that a term in a theory picks out some causal origin for a given set of observed phenomena. But if the theory does not do any work, as it were, we do not seem to have any means to prevent the term from picking out multiple entities—in general, we indeed need a reasonable amount of information in order to uniquely circumscribe the referent of a term. Inversely, if successive theories are not linked to the causal origin of the phenomena they are supposed to account for, we do not seem to have any means to ensure that they talk about the same entity if they say different things about it. Psillos's account can be seen as avoiding these two issues as follows: the term is first linked by causal contact to a set of referents $\{x_1, ..., x_n\}$, with $n \geq 1$; the core causal-explanatory description either fails to pick out any of these referents or selects a subset of them $\{x_1, ..., x_m\}$, with $m \leq n$; the tracking condition separates the remaining candidates (if any) according to whether they play a crucial explanatory role and ensures, at least in principle, that there is only one referent left in $\{x_1, ..., x_m\}$ at the end.

Consider, for instance, 'heaviness' in Galileo's later works (1632/1967; 1638/1974). Galileo takes heaviness to be a coarse-grained quality of bodies without committing to its deep nature or metaphysical status, and he uses the term to pick out some causal origin for the observed free fall of terrestrial bodies.[5] Newton's gravitational force does satisfy Galileo's phenomenological descriptions (e.g., heaviness is responsible for the differences in velocities between the earlier and later times of the free fall of a terrestrial body). Yet, on the face of it, Newton's gravitational force does not possess the kind-constitutive properties Galileo attributes to gravity in his later works (e.g., internal quality of terrestrial bodies). Hence, according to Psillos's account, 'heaviness' does not refer to Newton's gravitational force.

Turning now to the downsides of Psillos's account, note, first, that his appeal to kind-constitutive properties in (1999) is probably too restrictive. In particular, a given set of explanatorily indispensable properties might not constitute a well-delineated natural kind. Psillos (2012, p. 226) suggests using "stable identifying properties" instead, provided that they take part in the causal explanation of the observed phenomena with respect to the theory of interest. He does not, however, provide much detail about the notion

---

[5] I follow Koyré's interpretation of Galileo's notion of (absolute) gravity or heaviness in the *Discourses* and the *Dialogue* as a macroscopic, internal, and universal property of terrestrial bodies here (Koyré, 1966/1978, Part III, sec. 3).

of “stability” and the most immediate ways of specifying it do not seem to work. In particular, appealing to properties playing an indispensable role in the causal explanation of ϕ does not seem to be a reliable way of identifying stable properties (cf. Saatsi et al., 2009, pp. 365-6). There is indeed some significant leeway as to what constitutes a good enough explanation and thus as to whether some property is really needed to account for a given phenomenon. For instance, we might realize over time, perhaps thanks to a new theory, that we could have made the same predictions by bracketing some seemingly indispensable first step in their original derivation and starting directly from a less metaphysically loaded set of assumptions.

Now, even if we turn a blind eye to the issue of stability, Psillos’s account still appears to make referential success both too easy and too hard to achieve. His own example of Maxwell’s luminiferous ether and the classical electromagnetic field illustrates well the first case. Even if we accept that these putative entities play the same causal role (e.g., dynamical structure for the propagation of light waves at finite velocity) and share a set of “stable” core causal properties (e.g., continuous medium, repository of the kinetic energy of light), this does not seem to be sufficient for granting that ‘luminiferous ether’ refers to the classical electromagnetic field. As French (2014, pp. 4-5, 125) rightly points out in my view, some of the core individuating properties of Maxwell’s luminiferous ether, such as its mechanical nature and its molecular constitution, are not shared by the classical electromagnetic field. If we eliminate these properties as parts of what fixes the reference of ‘luminiferous ether’, referential success is achieved at the expense of replacing (as it were) the ether as a self-standing entity with clear identity conditions by a small cluster of stable properties. In this case, however, we face again the issue of radical referential indeterminacy, i.e., the issue that ‘luminiferous ether’ might refer to radically different types of entities, including a background space-time containing collections of photons and other sorts of particles.

Psillos’s account also seems to make referential success too hard to achieve: the same theoretical term might be associated with radically incompatible core causal descriptions in two different theories and still refer in both cases if these descriptions are used at different levels and in different circumstances. To give one striking example: (i) the gravitational force in classical Newtonian mechanics plays the same causal role as the curvature of space-time in classical General Relativity with respect to terrestrial gravitational phenomena; (ii) the term ‘gravity’ is usually associated with radically incompatible core causal-explanatory descriptions in the two theories (e.g., the gravitational force is non-local while gravitational effects propagate locally in standard curved space-times); (iii) and yet, near the Earth and more generally in restricted physical contexts where the curvature of space-time is sufficiently small and the characteristic time scale of the system is sufficiently large, we seem to be justified in identifying the causal origin of gravitational effects with a concrete variable relation between distant massive bodies.

These last two issues signal that Psillos's account of reference does not have appropriate resources to address what might be called the "problem of referential tracking" (which is a close cousin of the qua problem): namely, how should we restrict the set of putative referents of a term introduced to account for some observed phenomena given equally plausible candidates in the causal structure (or causal chain) underlying those phenomena? Suppose, for instance, that we want to identify the referent of 'heaviness' in Galileo's later account of gravitation. Should we restrict our focus to medium-size entities close to the surface of the Earth, bracket the Earth itself, and take the set of macroscopic terrestrial properties to be the appropriate locus of reference for the term? Should we zoom out, include the Earth in the domain of reference, and take 'heaviness' to refer to a force relating massive bodies? Should we zoom out even more to include sufficiently massive entities and take 'heaviness' to refer to the smoothly curved structure of space-time? Or perhaps should we rather zoom in, eliminate overly coarse-grained items, and take 'heaviness' to refer to collections of gravitons? Psillos's account, in other words, does not seem to be able to circumscribe some appropriate causal agent(s) given equally plausible candidates specified in more or less comprehensive physical contexts across scales. Let us see, then, how to address this issue while preserving Psillos's insights about the shared burden of reference fixing via causal contact and via satisfaction of a description.

3 Reference fixing and scale-dependence

Despite its deficiencies, Psillos's account has the merit of shifting the original problem of referential failure to the more tractable issues of referential tracking and stability. I will focus on referential tracking in this section, using ideas from the framework of effective theories. The solution, in a nutshell, is to identify the limited range of parameters over which the theory is empirically accurate and select the terms that refer to unobservable entities specifiable within the domain delineated by this range. I will discuss how the resulting account of reference fares with respect to the issue of referential stability in sections 4-5.

As we have seen, Psillos's appeal to "kind-constitutive" or "stable identifying" properties is far from ensuring that a term picks out some appropriate causal agent(s) at the relevant scale. We might be tempted to appeal to scientists' intention to refer to a particular entity. But this does not seem to be a good solution too. The scientists in question still need to extract information from a given theory in order to circumscribe potential candidates and select a specific target in their mind (including basic information about whether there is only one or several causal agents responsible for a given phenomenon). By following this route, we also take the risk of making the selection process overly sensitive to scientists' idiosyncratic beliefs and willingness to speculate about their subject matter. To avoid this, we might be tempted to appeal instead to the seemingly more objective "domain of applicability" of the theory of interest, that is, to the set of items—entities, properties, relations, and so on—specified by interpreting the theory literally. But this does not seem to work either. After all, specifying the domain of a theory does not only depend on its internal principles and constraints but also on how we intend to define its scope in the first place—and this, again, depends sensitively on particular interpretative choices. For an empiricist of a radical sort, for instance, this domain would reduce to a mere domain of terrestrial phenomena.

How can we constrain referential tracking without relying on scientists' particular beliefs? The framework of effective theories proves to be instructive here. Quite remarkably, once equipped with appropriate empirical data, the structure of an effective theory is such that it allows us to estimate in advance where its predictions are likely to break down and delineate its domain accordingly (I will provide more detail in section 4). The suggestion, then, is to restrict the set of putative referents of a term by means of the theoretical constraints obtained from the empirical limitations of the theory in which it appears:

Semantic Constraint: A term t in a theory T refers to some entity x only if x is specifiable within the domain of entities delineated by the range of parameters over which T is empirically accurate.

Let me illustrate this with two preliminary toy examples. (i) Suppose that a theory describes some lattice composed of elementary blocks of characteristic size a, such that both its theoretical descriptions and the predictions derived out of them depend on a. If those predictions start to become inaccurate, say, for $a \leq 10^{-6}$ m, the semantic constraint implies that the terms of the theory do not refer to elementary blocks with a microscopic or smaller characteristic size. (ii) Suppose that a theory describes a set of entities whose collective structure depends on the temperature T of the environment in which they evolve. If the predictions of the theory start to become inaccurate at very low temperatures, the semantic constraint implies that the terms of interest do not refer, say, to compounds made out of such entities and that exist only in such extreme physical circumstances. In each case, the semantic constraint offers a solution to the problem of referential tracking in the sense that it restricts the set of candidate referents for a term according to whether they fall within the empirical reach of the theory that purports to describe them.

To unpack further the content of this semantic constraint, let us have a look at a more realistic particle physics example (I will come back to gravitational cases in section 4). Consider Rutherford, Geiger, and Marsden's discovery of the atomic nucleus in the early 1910s after having bombarded thin metal foils with positively charged particles (see, e.g., Heilbron, 1968; Kragh, 2012, for more historical details). Rutherford accounted for the unexpected pattern of particle scintillations around the target by assuming that atoms have a nucleus, i.e., a highly compact (positive) charge distribution at their center. Energy and momentum conservation laws indeed implied that the mass of single electrons was too small for them to have a significant impact. Incoming particles had to interact with a sufficiently massive centered charge distribution for a sufficient number of them to be reflected in a symmetric manner. Their kinetic energy could even be used together with conservation laws to provide a lower bound on the mass and a higher bound on the effective radius of the charge distribution. More generally, the empirical success of Rutherford's theoretical model over a particular range of experimental conditions provided stringent constraints on the range of parameters involved in the description of the putative referents for 'atomic nucleus' (e.g., high mass, small radius, central atomic location, absolute charge varying with the metal considered). But without more discriminating experimental means, say, incoming particles with a sufficiently high energy, it was impossible to probe further the properties of these putative referents at shorter distances (e.g., the internal structure of complex atomic nuclei in terms of a collection of protons and neutrons or a set of quark and gluon fields in a localized configuration) or gain clear empirical evidence that the classical Newtonian and electromagnetic theoretical constraints involved in Rutherford's theoretical model would start to become inadequate at sufficiently short distances and high velocities.

This more realistic example raises three questions: (i) What does it mean, exactly, for some entity x to be "specifiable" within a given domain? (ii) What does it mean for some domain to be "delineated" by a limited range of parameters? (iii) What is involved in the determination of such a range?

Regarding (i)-(ii), I take it that an entity x is specifiable within a domain D if and only if x has at least all the properties required to belong to D (e.g., some appropriate characteristic size for the domain of atomic entities). We may increase or reduce the content of D—or, more precisely, consider more or less comprehensive domains—by deleting or adding properties on the list of entry requirements. We may also represent a quantifiable property by means of a parameter (or a variable) and further delineate D by relaxing existing constraints or fixing new limits on the values of this parameter (e.g., the domain of entities with a characteristic size larger than 10^{-11} m, which includes both atoms and larger entities but not atomic nuclei). In each case, we can think of the domain restriction in terms of a minimal or maximal resolution scale filtering out entities whose properties do not fit together within the set of constraints imposed on the domain. But we should be careful not to reduce the notion of scale to sizes or distance scales. In Rutherford's case, for instance, the restriction to high mass density scales excludes insufficiently compact charge distributions.

Regarding (iii), I am inclined to interpret the range of parameters used to delineate the domain of a theory in a qualified empiricist sense, i.e., as being directly determined by means of experimental procedures and previously tested theoretical relations. At Rutherford's time, for instance, the kinetic energy of incoming particles was already typically determined beforehand through their magnetic and electric deflections in a test material together with standard kinematical and electromagnetic theoretical relations (e.g., energy conservation, electrostatic potential energy). A similar set of theoretical relations was also required to determine a higher bound for the effective radius of the nucleus given the kinetic energy of incoming particles. Note that the extent to which a parameter is more or less theoretical is directly related to the set of theoretical constraints required to determine its values. The range of a more experimental parameter is also usually used to determine the range of a more theoretical parameter delineating the domain of the theory. We should thus be careful again not to restrict the notion of parameter to that of measurement parameter (or measurement scale).

The motivation for staying as closely empiricist as possible when assigning physical meaning to such parameters is twofold: (a) to ensure that the restrictions imposed on the domain of the theory are largely independent of its interpretation and thus insensitive to interpretative disagreements about what the world is made of in this domain (e.g., different kinds of particles in Rutherford's case); (b) to ensure that even the most theoretical parameters are anchored to experimental and observational facts (e.g., characteristic sizes, particle rest masses) and avail oneself of an objective basis for drawing boundaries between different domains. Of course, if needed, we may attribute further physical meaning to such parameters (or refine their interpretation) once we enter into the business of interpreting the content of the theory (e.g., definite vs. average location, physical vs. bare mass).

Now that the content of the semantic constraint is clarified, three comments are in order. First, if we wish to assess at a given time whether a term refers, we first need to look at the empirical success of the theory in which it appears. In general, this means restricting oneself to the range over which the theory has been found to be empirically accurate (as in Rutherford's case). In some cases, however, we may be able to estimate the limited range over which a theory is likely to remain empirically accurate even if we have not yet probed phenomena at the relevant scales in experiments or through observations (cf. section 4). Either way, the important point for now is that it is our assessment of referential success and not the semantic constraint itself that depends on the empirical evidence that we have for a theory at a given time.

Second, the semantic constraint fleshes out scientists' implicit scale-dependent referential practices; but it does not make the assessment of referential success depend significantly on scientists' and interpreters' intentions and expectations. We certainly need to pick a reasonable standard of measurement accuracy and the extent of the range over which the theory has been found to be empirically accurate depends on the experimental and observational achievements reached at a given time. But apart from that, we only need to assume that the predictions derived from the theory depend on, or can be associated with, a set of parameters and that, as we vary them, the comparison of predictions with empirical data determines the limited range over which the theory is empirically accurate.

Third, the semantic constraint appears to be good enough to rule out plausible cases of radical referential indeterminacy across scales. In Rutherford's case, the term 'atomic nucleus' refers to sufficiently coarse-grained and dense parts of matter. But it does not involve reference to individual neutrons for instance, or anything more fine-grained for that matter.[6] Likewise, in the Galilean case, if we restrict ourselves to the free fall of a body close to the ground and ignore the Earth, we automatically exclude any kind of instantaneous force between the Earth and the body within the set of putative referents for 'heaviness'. I will discuss in greater detail the issue of radical ontological discontinuities in sections 4-5. For now, what matters is that the new semantic constraint appears to be strong enough to prevent a term from referring to radically different causal agents in widely different domains (assuming, of course, that we are realists about sufficiently robust non-fundamental domains, as already emphasized in the introduction).

The next step is to implement the new semantic constraint within a general account specifying both the mechanism of reference fixing and the conditions under which a term refers. I will follow Psillos's lead here, starting first with the causal link condition (C). As we have seen, this condition is meant to ensure that a term in a theory is linked to some underlying causal agent given some observed phenomena. If we apply the new semantic constraint to (C), the set of referents picked out by causal contact is restricted to those that belong to the limited domain delineated by the range of parameters over which the theory is empirically accurate. The resulting account is thus context-dependent in an objective sense, i.e., reference fixing is indexed to particular levels and circumstances (or, more precisely, to some limited domain).

---

[6] The case of protons is more ambiguous. It arguably became only justified to believe that the extension of the term 'atomic nucleus' in Rutherford's model includes individual protons only after he obtained hydrogen nuclei by bombarding a nitrogen gas with alpha particles in 1917 (see, e.g., Kragh, 2012, sec. 6, for more historical details).

Consider now conditions (S) and (T). The main difficulty here is to separate the descriptions of the theory according to its empirical limitations. Again, the framework of effective theories proves to be instructive. The descriptions $D(\Lambda_1, ..., \Lambda_n)$ of an effective theory explicitly depend on a set of limiting parameters $\Lambda_1, ..., \Lambda_n$ ("cut-off scales") that are used to specify its limited range of empirical validity $\mathcal{R}$, say, a short-distance cut-off scale r_0 in the simple example of an effective theory whose predictions break down at short distances r. In typical cases, the structure of an effective theory is also such that we can separate these descriptions into two sets according to whether they provide reliable information within $\mathcal{R}$, say, into $\{D(r, r_0), r > r_0\}$ and $\{D(r, r_0), r \leq r_0\}$ in the simple example above. I will provide more detail below. For now, this suggests (together with the examples used so far) that we should restrict the set of referents picked out by causal contact to those satisfying the core causal-explanatory descriptions that are indexed to, or associated with, the limited range over which the theory at stake is empirically accurate.

Putting all of this together, the resulting account of reference CST* takes the following form. A term t in a theory T refers to an entity x under three conditions:

(C*) Causal link: t is introduced (or used) to pick out some causal origin x for a set of observed phenomena ϕ within the domain delineated by the limited range of parameters $\mathcal{R}$ over which T is empirically accurate;

(S*) Satisfaction link: x satisfies the core causal-explanatory description of ϕ associated with t in T;

(T*) Tracking: The core causal-explanatory description of ϕ is indexed to the limited range $\mathcal{R}$.

Consider again Rutherford's case. The term 'atomic nucleus' (or 'central atomic charge') is introduced to pick out some causal origin for the significant amount of scintillations observed at large scattering angles. Independently of the specific details of Rutherford's core causal-explanatory description, the set of putative referents is restricted to the domain where his theoretical model is empirically accurate (e.g., sufficiently compact and coarse-grained entities within atoms given the kinetic energy of incoming particles and previously tested theoretical relations). Rutherford's core causal-explanatory description in terms of a very small and massive positive charge distribution centered in each atom and responsible for the deflection of incoming positively charged particles via electromagnetic interactions restricts this set of putative referents. Note that this description is indexed to the range of parameters over which Rutherford's theoretical model is empirically accurate. By contrast, the description of the nucleus as a point particle plays no reference-fixing role (if it is to be taken seriously at all). Note as well that in Rutherford's case, the domain restriction in (C^*) is essential. Collections of more fine-grained entities, say, of protons and neutrons, do satisfy Rutherford's rather coarse-grained core causal-explanatory description of atomic nuclei. But by imposing condition (C^*), the putative referents of 'atomic nucleus' are constrained to lie within a sufficiently coarse-grained domain that excludes individual neutrons for instance.

I will explain how CST^* works with other physical examples in sections 4-5. For now, two remarks are in order. (i) CST^* is independent of issues regarding the choice and interpretation of the core causal-explanatory description. The association of a term with a particular set of interpreted descriptions should be seen as an input and CST^* as offering a verdict given this input. Otherwise, for the canonical examples discussed here, I assume for simplicity that the referential success of a term is assessed by taking the standard interpretation of its core causal-explanatory description (e.g., a highly compact positive charge distribution at the center of each atom in Rutherford's case). (ii) The story is likely to become more complicated in less fundamental areas of physics and the special sciences. Typically, in such cases, we need to specify a larger set of parameters and distinguish between the target of interest and background entities constituting some larger domain in which the target is included. Even in the Galilean case, for instance, we need to assume that the heavy body of interest is falling near the ground and look for putative referents of 'heaviness' within this physical context. In less mathematized scientific areas, we probably also need a more liberal notion of parameter to make room for non-quantifiable determinable properties with different determinate aspects. These adjustments, however, do not seem to pose any irreducible threat to restricting the domain of a theory according to its empirical limitations. As we will see below, we may restrict the domain of interest in the case of 'phlogiston' by appealing to a limited set of chemical substances involved in a limited set of possible reactions in addition to various parameters with well-delineated ranges, such as the weight and the volume of the substances involved.

4 The case of effective theories

We have seen that we can select terms in a theory according to whether they pick out referents specifiable within the domain of unobservable entities delineated by its empirical limitations. In this section, I will argue that there are good reasons to believe that the terms selected are referentially stable under theory change in the case of effective theories.

Consider first the standard Galilean and Newtonian laws of free fall, rewritten in their mathematically most simple modern formulation for conceptual clarity (see Table 1 below). The target system in the Galilean case is a heavy body dropped at some height $z(t)$ from the ground. The target system in the Newtonian case is a body of mass m located at some distance $r(t)$ from the center of the Earth, with M its mass and G the universal gravitational constant.

Galilean	**Newtonian**
$\frac{d^2z}{dt^2} = -g$ $g = \text{constant}$	$m\frac{d^2r}{dt^2} = -mg(r)$ $g(r) = \frac{GM}{r^2}$
g: heaviness of matter in a vacuum (universal quality of terrestrial bodies, local internal action).	$g(r)$: interaction force exerted by the Earth on the body per unit mass (relational property, action at a distance)

Table 1: The Galilean and Newtonian laws of free fall.

We could directly appeal to Newton's gravitational theory (resp. classical General Relativity) to evaluate the referential success of the term 'gravity' in the Galilean (resp. Newtonian) law of free fall. But suppose for the sake of the argument that we are dealing with each law at the time it is still a live concern and that we do not yet know anything about their respective successor. We can construct effective laws as follows. (i) Identify the limited range of parameters (or variables) beyond which the original law might become unreliable. For instance, we may have already found that the Galilean law makes slightly inaccurate predictions for heavy bodies dropped too far from the ground. Or we may suspect from the infinite value of $g(r) = GM/r^2$ in the limit $r \to 0$ that the Newtonian law is mathematically inadequate for describing gravitational effects between arbitrarily small bodies at very short distances (and, more generally, very strong gravitational effects).[7] (ii) Restrict explicitly this range by introducing some arbitrary cut-off scale: namely, a large-distance scale z_0 in the Galilean case and a short-distance scale r_0 in the Newtonian case. (iii) Include all the possible

---

[7] Of course, in the toy example of terrestrial bodies, we should ultimately restrict the range of r to distances larger than the radius of the Earth. Yet, as we will see below, extracting more general limitations from the asymptotic mathematical pattern of this law still proves to be relevant even in this specific context.

mathematical terms that depend on the cut-off scale and the parameter (or variable) at stake and that are allowed by the principles of the law, including its symmetry principles, with one arbitrary coefficient for each new term:

$$\text{For } z \ll z_0, \qquad \frac{d^2 z}{dt^2} = -g\Big(1 + a_1 \frac{z}{z_0} + a_2 (\frac{z}{z_0})^2 + ...\Big); \tag{1}$$

$$\text{For } r \gg r_0, \qquad m\frac{d^2 r}{dt^2} = -\frac{GMm}{r^2}\Big(1 + b_1 \frac{r_0}{r} + b_2 (\frac{r_0}{r})^2 + ...\Big), \tag{2}$$

with a_i and b_i arbitrary coefficients.[8] As I will explain below, those sorts of mathematical expansions allow us to encode the effects of new physics (if any) and assess its relative importance across scales.

Of course, in practice, the situation is more complicated than what these toy models may suggest. (i) Restricting the range of an effective theory is usually a highly non-trivial task. However, it does not require knowing the exact value or the underlying meaning of the cut-off scale (e.g., that z_0 is the radius of the Earth R_E and r_0 its Schwarzschild radius $2GM/c^2$, with c the speed of light). (ii) We often need more than one effective theory to cover the variety of phenomena related to a given domain. Note, however, that the theoretical landscape and the set of relevant parameters become simpler when we consider more fundamental theories (e.g., the bulk of the empirical limitations of the Standard Model of particle physics is likely to be captured by a single energy parameter).

To show how CST* works with these toy models, I first need to make explicit three of their most distinctive features, using the effective Newtonian law as my main example here.

First, the arbitrary cut-off scale in the effective law stands for its maximal predictive limit along a particular range of parameters. This interpretation is far from being straightforward (see, e.g., Rivat, 2021, for more detail). But it is motivated by the fact if we try to make the effective law as empirically accurate as possible at large distances $r > r_0$ by adding increasingly many higher-order terms $b_i(r_0/r)^i$ and fixing their parameter with new empirical data, its predictions become increasingly large close to r_0 and ultimately break down at this scale if we add infinitely many sufficiently significant terms in the expansion $\sum_i b_i(r_0/r)^i$.

Second, if we have appropriate data in some accessible regime, we can usually obtain a first estimate of the value of the cut-off scale r_0 by assuming

---

[8] Note that, in some cases, existing empirical data (or some other reason) may require us to drop some of these principles (e.g., finding that the Galilean law does not work so well at high altitudes requires us to break translation invariance along z).

that the coefficients are of order 1 (e.g., $b_i = O(1)$ in Eq. 2 above). We can even usually obtain a remarkably reliable estimate if we succeed in gleaning some information about the new physics, say, about the strength with which it couples to the known physics.

Third, and as already mentioned in section 3, we can use the cut-off scale r_0 to separate the descriptions $D(r, r_0)$ of the effective law into two parts, i.e., $\{D(r, r_0), r > r_0\}$ and $\{D(r, r_0), r \leq r_0\}$, and rely on the relative importance of their contribution to predictions to further approximate $\{D(r, r_0), r \gg r_0\}$ with the first few lowest-order terms in Eq. 2 for $r \gg r_0$. We may also further refine the range of parameters over which the predictions of the law are likely to remain empirically accurate, say, along velocity and mass density scales, and separate its descriptions accordingly.

I should emphasize that the structure of the original Newtonian law of free fall in Table 1 does not give us any such means of specifying its limited range of empirical validity and delineating its domain accordingly. If we take this law at face value and make too much of its remarkable empirical success at large distances at a given time, we might wrongly take it to offer a complete account of gravitation and describe some fundamental kind of entity with core properties specified, say, at arbitrarily short distance scales. Of course, cautious as we are, we may already have found good reasons to believe that the law is unreliable in such regimes, just like before. But even in this case, the structure of the law does not provide much information about the scales at which it is likely to become unreliable beyond the range over which it has been found to be empirically accurate at a given time.

The methods of effective theories, by contrast, allow us to parametrize our ignorance about its scope and identify precisely the set of descriptions that we have no reason to trust at the time it is still a live concern. If we have appropriate empirical inputs, these methods even allow us to estimate the range over which the effective law is likely to remain empirically accurate and restrict the putative referents of its terms accordingly without knowing anything (or much) about the next theory (if any). And if we have such an estimate and decide to apply CST* to the effective law at a given time, we will find that the term 'gravity' (or 'gravitational force') refers to a concrete variable relation at large distance scales between slowly moving entities with low mass density, and whose magnitude is given by the first few lowest-order terms in Eq. 2 for $r \gg r_0$.[9]

This is, of course, not enough to settle the issue of referential stability. So far, CST* only allows us to reduce the risk of referential failure. The structure of the effective Newtonian law indeed delineates the scales where it is likely to make inaccurate predictions and thereby provide false information

---

[9] Note that this effective interpretation does not require us to specify any kind of underlying local or non-local causal gravitational mechanism beyond the concrete variable relation between distant massive bodies.

about its target system. Thus, the law gives us good reasons to believe that its descriptions ranging over these scales are false and fail to be satisfied by anything real. And if we impose the tracking condition (T^*), we have a direct way to ensure that these descriptions do not play any reference-fixing role, i.e., to reduce the risk that the terms selected fail to refer to any of the candidates picked out by causal contact as specified by condition (C^*).

As it turns out, there are also good reasons to believe that the terms selected through CST^* are referentially stable in the case of effective theories. But to make this point, I need another one of their distinctive features: namely, that the most relevant descriptions for predictions in the range where the effective theory is empirically accurate are typically largely independent of descriptions that are relevant well beyond this range, whether these descriptions come from the effective theory itself or a new theory.[10]

To illustrate this, consider again the effective Newtonian law with only a few lowest-order terms in $1/r$. But suppose this time that we have discovered a radically new theory revealing that the predictions of the effective law are slightly inaccurate for $r \gg r_0$. The good news is that we can compensate for predictive discrepancies by adding higher-order terms and fixing the value of their coefficients with empirical inputs at large distances. Granted: the discovery of a new and seemingly better theory makes this move look somewhat ad hoc. Yet it shows that these higher-order terms encode the contributions of new physics at large distances according to their relevance and thus that these terms do not correspond to arbitrary modifications of the Newtonian law with no physical significance whatsoever. The ability of higher-order terms to stand for fine-grained features of new physics is also supported by explicit derivations of effective theories from more comprehensive theories (see, e.g., Donoghue, 1997, sec. 8-9, for the expression of the first-order relativistic and quantum corrections to Newtonian gravitation). And, in general, the structure of an effective theory is such that we can parametrize the contributions of any type of new physics at large distances up to an arbitrarily high degree of precision by adding increasingly many terms depending only on the degrees of freedom of the original theory and consistent with most, if not all, of its principles. That is, an effective theory is typically able to accommodate the effects of new physics within its restricted domain in its very own terms.

Now, the crucial point is that the contributions of higher-order terms become increasingly negligible at large distances $r \gg r_0$, no matter what the new physics looks like. And insofar as these higher-order terms stand for fine-grained features of new physics, this shows that the relevant descriptions

---

[10] Note that this feature needs to be distinguished from the more restricted type of independence from renormalization artifacts (see, e.g., Rivat, 2021, for more detail). I will also not discuss here the related issue of referential stability arising from renormalization-scale-dependent terms (e.g., 'electromagnetic coupling').

of the effective theory at large distances are largely independent of the details of this new physics. That is, these descriptions are largely independent of a large variety of more comprehensive and alternative accounts, which specify, in particular, the properties of this new physics at short distances. Hence, in the case of effective theories at least, the tracking condition (T*) selects core properties at a particular level, which, in general, do not significantly depend on more fine-grained (or coarse-grained) features. This, in turn, gives good reasons to believe that the terms of an effective theory selected through CST* pick out entities in a limited domain, independently of how a future theory will describe them within a more comprehensive domain.[11]

We might still worry about the threat of radical ontological discontinuities. For instance, Galileo's late description of gravity as an intrinsic coarse-grained quality of bodies is specifiable within the limited domain delineated by the empirical limitations of his law and largely insensitive to more fine-grained and coarse-grained features of the world. The same goes for Newton's description of gravity as a concrete variable relation between massive bodies. We thus seem to have good reasons to believe that the terms 'heaviness' in Galileo's law and 'gravitational force' in Newton's law are both referentially stable. And yet, on the face of it, their core causal-explanatory descriptions appear to be incompatible with each other, which seems to force us to conclude that 'heaviness' fails to refer (given that Newton's law is more successful than Galileo's).

In response, the first thing to note is that their respective core descriptions are indexed to distinct ranges of parameters. On a literal construal, these descriptions have their truth values fixed relatively to overlapping yet distinct limited domains and thus do not contradict each other, strictly speaking. To be sure, we may restrict Newton's core description to the range over which Galileo's law is empirically accurate and compare the two descriptions with respect to the same domain. There is no issue of incommensurability here. But insofar as we are dealing with a limited domain where we have ignored (or "integrated out") the Earth, we are forced to reinterpret the Newtonian law (including 'M' and 'R_E'). In this case, it does seem plausible to reinterpret the gravitational force per unit mass $g(z) = GM/(z + R_E) \sim GM/R_E = g$ as a macroscopic property of terrestrial bodies in the domain delineated by $z \ll R_E$. There are of course many metaphysical and formal details to fill in here, especially about the

---

[11] See Fraser (2018) and Williams (2019) for a similar argument about inter-scale insensitivity employing RG and EFT methods in the context of quantum field theory. See Ruetsche (2018; 2020; 2024); Rivat, (2019; 2021); Rosaler and Harlander (2019); Rivat and Grinbaum (2020); Bechtle et al., (2022); Koberinski and Fraser (2023); Dougherty (forthcoming), for a discussion. See also Ladyman and Lorenzetti (forthcoming), Robertson and Wilson (forthcoming), and Baron et al. (forthcoming) for discussions involving a broader notion of effective theory.

reduction of concrete relations to monadic properties in a limited domain and the metaphysical relations that hold between more or less coarse-grained domains.[12] This example is also far from ensuring that CST* does not fail in some cases. To assuage this worry, I will provide further examples below to support the claim that typical cases of referential failure concern domains that have not been put to the test yet and thus that CST* remains suitable.

5 Problematic historical cases

I will now briefly discuss how CST* works more generally in physics and show how it helps us to handle problematic cases in the history of physical sciences, such as 'luminiferous ether' and 'phlogiston'.

Note, first, that CST* allows us to reduce the risk of referential failure prospectively under relatively mild assumptions (see Vickers, 2017, sec. 4, for a similar suggestion). Suppose that we are able to identify the limited domain delineated by the range of parameters $\mathcal{R}$ over which the theory at stake is empirically accurate at a given time (e.g., the domain of sufficiently compact and coarse-grained subatomic entities in Rutherford's case) and separate its descriptions according to whether they characterize only this domain (e.g., the effective size and mass of a charge distribution vs. its fine-grained structure). In this case, we seem to have good reasons to take the descriptions ranging beyond $\mathcal{R}$ to be more likely to be false and thus more conducive to referential failure than the others (e.g., the nucleus as a point particle). Of course, in contrast to effective theories, the structure of the theory might not give us any information about its putative predictive failure in new regimes. We might thus think that it is more reasonable to remain agnostic about the referential success of its terms in the corresponding domains. We do not seem, however, to be justified in remaining equally agnostic about terms whose referents are specifiable within the domain where the theory has been found to be empirically accurate (all else being equal). CST* enjoins us to take the safe side of this referential prescription when we are in the business of assessing referential success: look first for unobservable entities that fall within the empirical reach of the theory and assume that its terms fail to refer to anything real in unexplored domains until there is evidence to the contrary.

---

[12] I suspect that part of the recent debate surrounding the alleged conflict between the posits of Newtonian mechanics and classical General Relativity comes from underlying disagreements about the existence of non-fundamental coarse-grained domains and the relations between more or less fundamental domains (e.g., Ruetsche, 2018, 2020; Egg, 2021; Saatsi, 2022). One should also be cautious about assessing the ontological merits of Newtonian gravitation by treating classical General Relativity as a putatively fundamental and complete theory (or bracketing the fact that both theories enjoy a wide variety of formulations and interpretations).

Suppose now that we are also able to separate different causal components entering into the explanation of phenomena according to $\mathcal{R}$ (e.g., the effective size and mass of a charge distribution vs. its underlying subatomic constituents with respect to the detection of reflected particles). Suppose furthermore that the core causal-explanatory descriptions indexed to $\mathcal{R}$ are largely independent of a large variety of competing and plausible causal-explanatory descriptions indexed to larger ranges (involving, e.g., assumptions about the subatomic structure of the atomic nucleus).[13]

In this case, CST* also seems to give us a reliable basis for identifying terms that are more likely to be referentially stable than others. First, as we have just seen, we seem to have good reasons to trust more the descriptions of causal components that enter into the explanation of phenomena at the level and under the circumstances where the theory is empirically accurate (compared to other descriptions). Second, the independence of these privileged descriptions with respect to a variety of competing ones characterizing larger domains seems to increase our confidence that this will also be the case with the new and more comprehensive theory (if any). Of course, as in the case of effective theories, there is no way to fully ensure that the descriptions of this new theory will not be radically incompatible with those of the "old" theory once they are restricted to its limited domain and thus that the old theory will not be fully thrown away as radical skeptics might envisage it. But if we are able to come up with plausible and more comprehensive causal-explanatory alternatives that do not significantly affect the privileged descriptions, we do seem to have better reasons to believe that they will remain largely unaffected under theory change than suspend our judgment on this matter. In terms of CST*, we seem to have better reasons to believe that a term in a theory is referentially stable than not if its putative referent falls within the empirical reach of the theory and does not significantly depend on a large variety of putative causal components specifiable within larger domains.

Turning now to problematic historical cases, consider first Maxwell's luminiferous ether (see, e.g., Schaffner, 1972; Darrigol, 2000, for more historical details). According to CST*, the term fails to refer because some of the core properties of the luminiferous ether, e.g., that the ether has a molecular structure with fine particles and that light waves are continuously transmitted by means of the mechanical action of this molecular structure, characterize domains where Maxwellian theories of the luminiferous ether are

---

[13] The idea of robustness across levels and circumstances at work in this last assumption is akin to the kind of autonomy discussed by Wimsatt (2007, chap. 10, esp. pp. 216-21) in the general case and by Williams (2019) in the particular case of high-energy physics. Note that Rutherford's core causal-explanatory description was robust under a variety of plausible assumptions concerning the atomic nucleus (e.g., point particle vs. spherical charge distribution, interaction pattern between the nucleus and atomic electrons).

empirically inaccurate (and thus where we have good reasons to take them to be false). Moreover, at the time these theories were still a live concern, scientists had no reason to take ‘luminiferous ether’ to refer either insofar as they did not have any good experimental access to short distance scales and thus any good evidence regarding the fine-grained structure of light. By contrast, there was strong evidence from observed diffraction and interference patterns that light had a wave-like structure at sufficiently large distance scales. This large-distance wave-like structure was also robust across the different kinds of luminiferous media posited by physicists at the time. It was thus reasonable to think that one had reliable epistemic access to, and thus speak about, light waves and oscillating wave-like patterns (which explains the empirical and explanatory success of 19th-century theories of the luminiferous ether). But contra Psillos, there was not any good reason and there is still not any by our current light to take ‘luminiferous ether’ in Maxwellian theories to refer to anything real, and even less so to the classical electromagnetic field.

Consider next the case of phlogiston (see, e.g., White, 1932; Siegfried, 2002, for more historical details). The core identifying properties of this entity include being contained within different types of substances, such as combustible ones, and released, in particular, during combustion and calcination processes. Now, at the time the phlogiston theories were still a going concern, namely, before the new oxygen theory of chemistry became increasingly popular by the end of the 18th century, the phlogiston theorists did not have any good independent experimental constraints to further specify the exchange process and the substance(s) exchanged during combustion and calcination. They had clear evidence that something was exchanged. They were even justified in speaking about a “principle of inflammability”, a “fiery principle”, or a “principle of combustion” if the corresponding terms were taken to refer to some chemical agent exchanged without any assumption as to whether it is emitted or absorbed (in the same way as we might speak about gravity without specifying its fine-grained propagation mechanism). But the evidence available at the time was not sufficient to believe that ‘phlogiston’ refers insofar as some of the core properties of its putative referent were specified at a too “fine-grained” level of the exchange process for which the phlogiston theory was not shown to be empirically accurate. For instance, it was common to assume that phlogiston was lighter than air in order to explain away the typical increase in weight of metals after calcination. The phlogiston theorists, however, did not have the experimental means of testing this assumption on independent grounds.

By studying a larger range of reactions involving different types of substances in a wider variety of well-controlled experimental situations, such as sulfur and phosphorus in closed vessels involving a limited amount of air and water, the advocates of the new oxygen theory could get a firmer experimental hold on the correct locus of the substance(s) exchanged during combustion, calcination, respiration, and other processes.[14] In particular, if one has good reasons to believe that the principle of conservation of weight applies universally to chemical reactions, there is strong evidence that a gain of weight in most metals after calcination arises because of the participation of an external substance. The range of parameters at work in this case is certainly much more complex and indefinite than in the case of more contemporary physical examples. But we can still restrict it by means of the different types, volumes, and weights of the substances involved in distinct chemical reactions, and assess whether the properties of the target fit together within these constraints.

6 Conclusion

I have argued that the apparent failures of reference over the course of the history of science are most reliably analyzed by focusing on the empirical limitations of theories. We should take a term in a theory to refer to some entity only if it can be specified within the limited domain of unobservable entities delineated by the range of parameters over which the theory is empirically accurate. If a term fails to refer according to this principle of selective reference, as it is the case for 'luminiferous ether' and 'phlogiston', we might still find that other terms in the theory pick out entities that fall within its empirical reach. We might even be able to gain some confidence about their referential stability if their description does not depend significantly on a large variety of plausible alternative descriptions characterizing more comprehensive domains. I have shown that the framework of effective theories provides us with a paradigmatic set-up for implementing this selective strategy successfully. If we cannot directly use this framework, it still provides us at least with a blueprint for assessing referential success in the case of physics. And in both cases, insofar as it is often, if not always, possible to determine at least partially the empirical limitations of theories before they are superseded and gain some confidence about the robustness of their descriptions within the corresponding domain, this strategy seems to provide us with a principled and reliable way of distinguishing between referential success and failure from the perspective of each theory.

Acknowledgments: I am very grateful to David Albert, Jenann Ismael, Philip Kitcher, Achille Varzi, David Wallace, and Porter Williams for very fruitful dis-

---

[14] This is, of course, not to say that all the terms of Lavoisier's theory are referentially successful (e.g., 'caloric') or that Lavoisier's theory was more empirically adequate than the phlogiston theory in every respect at that time. See Chang (2010) for a recent reevaluation of the Chemical Revolution and Blumenthal and Ladyman (2017; 2018) for a response.

cussions and/or feedback on earlier versions of this manuscript. I would also like to thank audiences in South Bend, Mons, Berlin, Munich, Paris, and Nice, as well as anonymous referees, for very helpful and stimulating exchanges. This work has been partly supported by the Volkswagen Foundation [grant no. 9D181].

References

Baron, S., B. Le Bihan, and J. Read (forthcoming). Scientific Theory and Possibility. *Erkenntnis*. Available at https://philsci-archive.pitt.edu/24581/.

Bechtle, P., C. Chall, M. King, M. Krämer, P. Mättig, and M. Stöltzner (2022). Bottoms Up: The Standard Model Effective Field Theory from a Model Perspective. *Studies in History and Philosophy of Science 92*, 129–143.

Blumenthal, G. and J. Ladyman (2017). The Development of Problems within the Phlogiston Theories, 1766-1791. *Foundations of Chemistry 19*(3), 241–280.

Blumenthal, G. and J. Ladyman (2018). Theory Comparison and Choice in Chemistry, 1766-1791. *Foundations of Chemistry 20*(3), 169–189.

Chakravartty, A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable.* Cambridge: Cambridge University Press.

Chang, H. (2010). The Hidden History of Phlogiston: How Philosophical Failure Can Generate Historiographical Refinement. *Hyle 16*(2), 47–79.

Darrigol, O. (2000). *Electrodynamics from Ampère to Einstein.* New York: Oxford University Press.

Donoghue, J. (1997). Introduction to the Effective Field Theory Description of Gravity. In F. Cornet and M. J. Herrero (Eds.), *Effective Theories*, pp. 217–240. World Scientific Publishing Company.

Dougherty, J. (forthcoming). Effective and Selective Realisms. *The British Journal for the Philosophy of Science*. Available at http://philsci-archive.pitt.edu/21830/.

Egg, M. (2021). Quantum ontology without speculation. *European Journal for Philosophy of Science 11*(1), 32.

Fraser, J. D. (2018). Renormalization and the Formulation of Scientific Realism. *Philosophy of Science 85*(5), 1164–1175.

Frege, G. (1892/1980). On Sense and Reference. In M. Black and P. Geach (Eds.), *Translations from the Philosophical Writings of Gottlob Frege* (3rd ed.)., pp. 42–55. Oxford: Blackwell.

French, S. (2014). *The Structure of the World: Metaphysics and Representation.* Oxford: Oxford University Press.

Galilei, G. (1632/1967). *Dialogue Concerning the Two Chief World Systems, Ptolemaic and Copernican* (2nd ed.). Berkeley: University of California Press.

Galilei, G. (1638/1974). *[Discourses on the] Two New Sciences.* Madison: University of Wisconsin Press.

Heilbron, J. L. (1968). The Scattering of α and β Particles and Rutherford's Atom. *Archive for History of Exact Sciences 4*(4), 247–307.

Hoefer, C. and G. Martí (2020). Realism, Reference & Perspective. *European Journal for Philosophy of Science 10*(3), 38.

Kitcher, P. (1993). *The Advancement of Science: Science Without Legend, Objectivity Without Illusions.* Oxford: Oxford University Press.

Koberinski, A. and D. Fraser (2023). Renormalization Group Methods and the Epistemology of Effective Field Theories. *Studies in History and Philosophy of Science 98*, 14–28.

Koyré, A. (1966/1978). *Galileo Studies.* Atlantic Highlands: Humanities Press.

Kragh, H. (2012). Rutherford, Radioactivity, and the Atomic Nucleus. Available at http://arxiv.org/abs/1202.0954.

Kripke, S. (1972). *Naming and Necessity.* Cambridge: Harvard University Press.

Kroon, F. (1985). Theoretical Terms and the Causal View of Reference. *Australasian Journal of Philosophy 63*(2), 143–166.

Ladyman, J. (2021). Structure Not Selection. In T. D. Lyons and P. Vickers (Eds.), *Contemporary Scientific Realism and the Challenge from the History of Science*, pp. 243–256. Oxford: Oxford University Press.

Ladyman, J. and L. Lorenzetti (forthcoming). Effective Ontic Structural Realism. *British Journal for the Philosophy of Science*. Available at https://philsci-archive.pitt.edu/22790/.

Levi, M. (2023). A Theory of Theories. *CERN Courier 63*(1), 37–40.

Psillos, S. (1999). *Scientific Realism : How Science Tracks Truth*. London: Routledge.

Psillos, S. (2012). Causal Descriptivism and the Reference of Theoretical Terms. In A. Raftopoulos and P. Machamer (Eds.), *Perception, Realism and the Problem of Reference*, pp. 212–238. Cambridge: Cambridge University Press.

Putnam, H. (1975). The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science 7*, 131–193.

Rivat, S. (2019). Renormalization Scrutinized. *Studies in History and Philosophy of Science Part B 68*, 23–39.

Rivat, S. (2021). Effective Theories and Infinite Idealizations: A Challenge for Scientific Realism. *Synthese 198*(12), 12107–12136.

Rivat, S. and A. Grinbaum (2020). Philosophical Foundations of Effective Field Theories. *The European Physical Journal A 56*(3), 90.

Robertson, K. and A. Wilson (forthcoming). Theoretical Relicts: Progress, Reduction, and Autonomy. *British Journal for the Philosophy of Science*. Available at https://orcid.org/0000-0002-0645-0761.

Rosaler, J. and R. Harlander (2019). Naturalness, Wilsonian Renormalization, and "Fundamental Parameters" in Quantum Field Theory. *Studies in History and Philosophy of Modern Physics 66*, 118–134.

Ruetsche, L. (2018). Renormalization Group Realism: The Ascent of Pessimism. *Philosophy of Science 85*(5), 1176–1189.

Ruetsche, L. (2020). Perturbing Realism. In S. French and J. Saatsi (Eds.), *Scientific Realism and the Quantum*, pp. 293–314. Oxford: Oxford University Pres.

Ruetsche, L. (2024). The Perils of Approximate Ontology. *Synthese 204*(4), 122.

Russell, B. (1905). On Denoting. *Mind 14*(56), 479–493.

Saatsi, J. (2005). Reconsidering the Fresnel-Maxwell Theory Shift: How the Realist Can Have Her Cake and EAT It Too. *Studies in History and Philosophy of Science Part A 36*(3), 509–538.

Saatsi, J. (2022). (In)effective Realism? *European Journal for Philosophy of Science 12*(2), 1–16.

Saatsi, J., S. Psillos, R. G. Winther, and K. Stanford (2009). Grasping At Realist Straws. *Metascience 18*(3), 355.

Schaffner, K. (1972). *Nineteenth-century Aether Theories.* New York: Pergamon Press.

Siegfried, R. (2002). *From Elements to Atoms: A History of Chemical Composition.* Philadelphia: Amer Philosophical Society.

Stanford, K. (2003a). No Refuge for Realism: Selective Confirmation and the History of Science. *Philosophy of Science 70*(5), 913–925.

Stanford, K. (2003b). Pyrrhic Victories for Scientific Realism. *Journal of Philosophy 100*(11), 553–572.

Stanford, K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives.* Oxford: Oxford University Press.

Stanford, K. (2015). "Atoms exist" is Probably True, and Other Facts That Should Not Comfort Scientific Realists. *Journal of Philosophy 112*(8), 397–416.

Vickers, P. (2017). Understanding the Selective Realist Defence Against the PMI. *Synthese 194*(9), 3221–3232.

Vickers, P. (2022). *Identifying Future-Proof Science.* Oxford: Oxford University Press.

Votsis, I. (2011). The Prospective Stance in Realism. *Philosophy of Science 78*(5), 1223–1234.

Weinberg, S. (1999). What is Quantum Field Theory and What Did We Think It Was. In T. Y. Cao (Ed.), *Conceptual Foundations of Quantum Field Theory*, pp. 241–251. Cambridge: Cambridge University Press.

White, J. H. (1932). *The History of the Phlogiston Theory.* London: E. Arnold & co.

Williams, P. (2019). Scientific Realism Made Effective. *The British Journal for the Philosophy of Science 70*(1), 209–237.

Wimsatt, W. (2007). *Re-Engineering Philosophy for Limited Beings.* Cambridge: Harvard University Press.

Worrall, J. (2020). Structural Realism: The Only Defensible Realist Game in Town? In W. J. Gonzalez (Ed.), *New Approaches to Scientific Realism*, pp. 169–205. De Gruyter.