Realism, Reference & Perspective

Carl Hoefer & Genoveva Martí
ICREA and University of Barcelona

Keywords: realism; scientific realism; reference; theoretical entities; unobservable entities; reference to unobservable entities; descriptivism; direct reference; causal-historical account of reference; holism; anti-realism

Abstract:

This paper continues the defense of a version of scientific realism, Tautological Scientific Realism (TSR), that rests on the claim that, excluding some areas of fundamental physics about which doubts are entirely justified, many areas of contemporary science cannot be coherently imagined to be false other than via postulation of radically skeptical scenarios, which are not relevant to the realism debate in philosophy of science. In this paper we discuss, specifically, the threats of meaning change and reference failure associated with the Kuhnian tradition, which depend on a descriptivist approach to meaning, and we argue that descriptivism is not the right account of the meaning and reference of theoretical terms. We suggest that an account along the lines of the causal-historical theory of reference provides a more faithful picture of how terms for unobservable theoretical entities and properties come to refer; we argue that this picture works particularly well for TSR. In the last section we discuss how our account raises concerns specifically for perspectival forms of SR.

1. Introduction

The purpose of this paper is to continue the defense of a novel form of scientific realism, first sketched in Hoefer (2019). Tentatively called "Tautological Scientific Realism" (TSR), the core ideas of this view are the following. (i) Scientific realism should exclude fundamental physics theories from its scope. (ii) The dramatic leap forward in many of the sciences in the 20th century justifies a realist stance for many *parts* of science. By contrast, despite all the wonderful episodes contained therein, the history of science before the 20th century is no place to try to extend or defend the claims of scientific realism (SR). (iii) Many areas of contemporary science have attained such a level of interconnection (between different domains of knowledge) and such a variety of sources of confirmation that it is now simply *no longer possible to coherently doubt the* *approximate truth of the core "lore" in those areas, without doing so by means of one or another radical-skeptical scenario.*[1] And we believe it is almost tautological to maintain that if you have what seems to be a great deal of evidence in favor of some piece of scientific lore, and if you cannot see any coherent way to doubt the at least approximate truth of it without resorting to postulating some radical-skeptical scenario (evil demons, brains in a vat, the world as computer simulation, etc.), then you should indeed believe it. Hence, TSR.

After giving a statement and some defense of the core ideas of TSR in section 2, we will go on in sections 3 and 4 to examine how to handle one of the great bugbears of scientific realism: the threats of meaning change and reference failure that come out of the Kuhnian tradition and which depend on a descriptivist understanding of how terms like 'phlogiston' and 'electrons' refer, when they do. We argue that descriptivism is not the right account of the semantics of scientific kind terms and property terms, and argue that an account along the lines of the causal-historical approach to reference, an account that fits TSR especially well, gives the correct picture. Thus, it is possible to show that the threat of our still being, unbeknownst to us, engaging in *empty talk about nothing* when we say things about electrons and viruses, is a mere chimera. In section 5 we will address how our arguments and our account of reference bears on current doctrines of perspectival scientific realism.

2. Tautological Scientific Realism

Since *scientific realism* became a named philosophical viewpoint and an area of debate in Anglo-American philosophy of science, in the 1970s, a number of powerful and persuasive arguments have been mounted against its most prominent forms. Why doesn't it just die out? Most philosophers have the experience, at one time or another, of recognizing the potency of anti-SR arguments and taking seriously the possibility that our best current science may be fundamentally wrong in many important ways, which will only be seen by future generations. But then we go back to our daily lives, watching TV and using the internet, getting results from blood tests, taking antiviral medications and so forth. And within days, we are once again thinking that scientific realism *must* be right, somehow, despite the opposing arguments.

1 We find the word 'lore' to be a convenient term for the bits of knowledge that we want to talk about in explicating TSR. 'Lore' has the right connotations: long-established and accepted things, but ones which perhaps should not be assumed to be true. By contrast, words like 'knowledge' and 'fact' carry the implication of truth, which is question-begging in the context of discussing and defending SR.

A nearly universal reaction to the forceful arguments against SR has been to go *selective:* to defend that some *parts* of our best science must be correct, or correct in a certain respect, even if truth-or-approximate-truth cannot be defended across the board. Structural realism urges that mature sciences get right certain structural features of the world. Psillos' selective SR maintains that the parts of science that were operative in making correct novel predictions should be held (to have been, and to now be) true-or-approximately-true. Entity realists say that we can believe at least in the existence of certain unobservable theoretical entities, although not the theories in which they figure, once certain conditions obtain.[2]

TSR is also a form of selective SR, but different from all the ones just mentioned. In the first place, TSR notes that the most convincing examples of scientific revolutions that completely change our ontology and ideology in science, since the turn of the 20th century, have been exclusively in the domain of fundamental physics theory.[3] Moreover, the foundations of current physics (quantum mechanics and quantum field theory, when taken at face value) are still incoherent, and most physicists as well as philosophers of physics expect that future revolutions will occur that radically change the ontology and ideology of the fundamental nature of matter and space or spacetime. This being the state of things, the path of wisdom is to exclude our current best fundamental physics theories from the domain of stuff that SR wishes to claim is true-or-approximately-true. Despite the amazing empirical successes of those theories, their contents - the theories, and to some extent at least the ontology - have to be put into a quarantine zone.[4] TSR maintains that it is possible to quarantine off the still-dubitable parts of fundamental physics from the rest of contemporary science, and then defend an appropriately qualified truth claim for vast areas of the rest of our current science.

The reason it is possible to defend the truth of many other areas of science has to do with another way in which TSR is selective. TSR does not try to claim that we can find aspects of putatively successful theories from earlier centuries and argue that they were both true and maintained into current-day science. Rather, TSR restricts its claim to the contents

2 For critical discussion of selective realism, see Tulodziecki (2017).
3 The Darwinian synthesis in biology, discovery of the mechanisms of heredity (DNA and RNA and other things), and some other major discoveries have occurred outside of physics, of course. But in these fields scientists were well aware beforehand that they did not have the truth or anything very close to it. So the discoveries did not have the character of radical overthrow of earlier beliefs that one sees in some physics revolutions.
4 We take the useful term 'quarantine' from Callender (2019). Callender is more pessimistic about the possibility of erecting a serviceable quarantine than we are. See Hoefer (2019) for more detailed arguments for the need to exclude fundamental physics from the domain of applicability of SR.

of what we *now* consider to be established, known scientific truths. Roughly speaking, the most important parts of what we now regard as established became clear and stabilized by the mid-late 20th century. In broad-brush terms, we can mention as part of this established lore some key parts of biology (basic facts about DNA and RNA and their role in cell reproduction, sexual reproduction, and evolution; numerous facts about bacteria and viruses and their role in illnesses; facts about how cells produce energy, how oxygen is transported from lungs to cells in multi-celled organisms, and so on). We can also mention central parts of chemistry, lore that gives us a grasp on how atoms bind into molecules, what happens in a huge range of chemical reactions, and so forth. We can also mention the core elements of electrical and electronic engineering, such as the knowledge that allowed us to make radios and amplifiers in the 1920s, transistors in the 1960s, and digital memory chips and processor chips in the late 20th century.[5]

Why is all of this lore not still uncertain, subject to refutation by future scientific revolutions? The key claim of TSR is this: once the varieties of evidence for, and the inter-entanglements between, bits of (putative) knowledge reach a certain richness and complexity, the only ways to genuinely conceive of the *serious* falseness of *important* parts of the accepted lore will all involve some sort of radical-skeptical scenario, such as Descartes' evil demon/deceiver, or Putnam's brain in a vat, or the more recent idea that we might be all parts of a hugely sophisticated computer simulation, or the apocalyptic worry that perhaps the laws of nature will suddenly disappear or change overnight. But radical skeptical scenarios put into doubt every bit of our knowledge, not just scientific knowledge. For the purposes of debates on SR in the context of philosophy of science, such scenarios can be legitimately set aside. And if it is really not possible to see how the best-established lore of (many parts of) current science can be coherently doubted without adverting to a radical skeptical scenario, then realism is victorious. This is what TSR claims to be, in fact, the case.

It is easy, we hope, to see why it is intractably hard to imagine some central features of modern science *outside* of physics, involving unobservable entities and processes, being

5 At this point some philosophers may object that we are just expressing our own subjective confidence in current science, and note that people in (say) the late 19th century might well have felt a similar confidence despite having an enormous number of beliefs that we now regard as false. But we believe that, despite the occasional expressions of extreme confidence that can be found in some statements of *some* prominent scientists, in fact many scientists back then understood that their theories were incomplete and quite possibly false in important respects. An example of this would be Darwin's acknowledgment that he did not know what the mechanisms of inheritance were. An exploration of how the confidence of scientists in the late 20th century and beyond compares with the confidence of earlier generations is a project beyond the scope of this paper.

substantially mistaken. Take for example our belief that viruses of many types exist, can be seen with high-powered microscopes, grown in lab cultures, and that they propagate "in the wild", causing various diseases in multi-celled hosts.[6] We have such an enormous variety of *epistemic handles* on viruses and their roles in causing diseases that to imagine their non-existence or that they exist but do not do any of the things we attribute to them ends up being just as hard as the fictional example Saul Kripke invented and used in *Naming and Necessity*: imagining that cats turn out to all be robots rather than living organic beings native to the Earth. To imagine our virus lore turning out to be substantially false requires setting up some scenario in which a dozen or more independent types of evidence about (and types of interaction we have with) viruses all turn out to be completely misleading. The putative story one tells for how this could be so will require adjustments in other pieces of accepted lore. For example, if we imagine trying to suppose that electron microscopes that we thought allowed us to image viruses systematically mislead us, we have to explain then why they *appear to* work so well for imaging and manipulating other things, like microscopic circuits we build for computation.[7] These changes or reinterpretations would presumably necessitate still other changes, in further areas of lore. In most cases, we suggest, a chain reaction would result that winds up forcing wholesale changes across most or all of our accepted lore; in effect, we would be pushed into a radical skeptical scenario of one kind or another as the only way to imagine carrying through the rejection of the initial piece of lore (that viruses with certain characteristics exist and cause diseases). If we set aside radical skeptical scenarios, such things are just not genuine epistemic possibilities. Thus, the things that TSR enjoins us to go ahead and trust to be true-or-approximately-true are just the sorts of things we mentioned above that feature in our daily lives and which return us to feeling that SR must somehow be right, after the spell of anti-realist arguments wears off.

6 Hoefer (2019) explores briefly the justification of TSR's claim in the context of a couple of different pieces of scientific lore outside of fundamental physics.

7 Note that the epistemic handles we have on a kind of entity such as viruses, are *not* simply bits of descriptive beliefs that we have about them (although our beliefs do of course play a role in how epistemic handles get established and how they relate to the things and facts that they are handles on). For example, *electron microscopes* – the things themselves and their behaviors, as well as our beliefs about them – constitute part of our epistemic handles on viruses. And the network of interactions involved in microscopy are also part of our epistemic handles on viruses, despite not all of them being *about* viruses in any direct sense. So epistemic handles include types of evidence about (or for) some thing or factual proposition; types of interaction with a thing; as well as ways of knowing and doing that are integrated into our web of beliefs and practices.

A couple of things are important to note about TSR. First, TSR is unlike most earlier forms of SR in not relying, at all, on *inference to the best explanation* (IBE) for its central argument. We think that the objections raised against IBE by van Fraassen and others are correct, and that SR must find a way to defend itself without relying on IBE. The same goes for the closely related "no miracles argument" first offered by Putnam, in so far as it is read as enjoining us to believe in the approximate truth of our best scientific theories *because such truth would be the best explanation for the success* of those theories. While the explanatory claim is perhaps correct, it is not enough to warrant belief in that approximate truth. But there is another aspect of Putnam's argument that lines up well with TSR: the idea that it would require positing a *miracle* - or as we prefer to say, positing some radical skeptical scenario - to coherently imagine that many core aspects of contemporary science's picture are not even approximately true.

Second, we note that in its central claim, TSR is essentially a direct denial of one of the more recent challenges to SR, namely Kyle Stanford's *problem of unconceived alternatives* (PUA) (Stanford (2006)). The PUA challenges SR by pointing out that, on top of the historically already-known examples of also-successful-but-false theories that the history of science (and contemporary science, in the domain of fundamental physics) gives us, SR must contend with the possibility that the *real* truth of things is actually entirely out of our cognitive grasp so far. How can we claim to be able to rule this out? And if we cannot, how can we sustain any significant form of SR?

When it comes to fundamental physics, TSR freely admits that Stanford is right. But at the same time, TSR insists that the content and domain of fundamental physics[8] is *unique* in this regard, compared to the other highly established areas of physical and biological science. Moreover, and crucially, no matter what the hypothesized future physics looks like, it is not going to undercut the correctness of the *coarse-grained* lore of physics - things like the existence of various types of stable elementary atoms, that they are mostly composed of electrons, protons and neutrons in certain stable bound configurations, and so on. At this point, given the enormous number of epistemic handles we have on these parts of physics lore, only a skeptical scenario could raise such doubts. Thus, high school physics texts are not going to be radically re-written; just, perhaps, their prefaces or final

[8] We include much of cosmology in the ambit of fundamental physics that must be quarantined for now, because its fundamental posits and models are strongly connected with fundamental physical theories such as General Relativity, the Standard Model of particle/field physics, and sometimes even speculative quantum gravity theories.

chapters. And the physics and chemistry that biologists need to learn in order to do their daily work will not have changed, for the most part at least, either.[9]

The qualification 'for the most part' is crucial, because of course we cannot rule out that future physics makes possible effects in biology or chemistry that are currently considered physically impossible. Perhaps cold fusion will turn out to be possible, or humans will be able to utilize EPR entanglement for telepathic communication faster than light. No matter; the scientific stories about what's going on in most everyday physical happenings, where we already have a reasonably complete story in place, will not be undermined by such novelties. TSR's claim is just that most of what scientists think they know with certainty, in large swaths of current science, is indeed true-or-approximately-true. TSR does not claim that current science already knows most of what there is to know.

Returning to PUA, we maintain that it can thus be quarantined as well: it is a serious problem only in the domain of fundamental physics. Now, in Stanford (2006) most of the discussion concerns unconceived alternatives in biology, especially the biology of inheritance of organismic traits. But crucially, Stanford's case studies come from the history of science, not from contemporary biology; and TSR makes no realist claim concerning the scientific theories of those earlier epochs. Indeed, the defender of TSR can happily concede that unconceived (radically different, and superior) alternatives have indeed been generically possible throughout the history of science prior to the mid-late 20th century. What she denies is that one can legitimately induce that they are *still* possible, for certain areas of contemporary science.

So far we have presented the main characteristics of TSR, and have motivated why we think it may be stronger than earlier forms of SR. Now we turn to the discussion of how TSR can overcome apparent challenges to the truth of scientific claims, particularly claims about the existence of unobservable theoretical entities: challenges that arise from considerations of reference and meaning.

3. The referential, existential threat to SR

9 Note that this is not because either chemistry or biology can be said to be domains in which material things obey classical Newtonian physics. In both sciences facts about ions, chemical bonds and a number of purely quantum effects are frequently relevant. But those relevant facts can be captured in terms of the entities that we believe to be "here to stay" and their behaviors, which we claim to be included in the safe areas of physics lore.

Many philosophers have thought that considerations of meaning and reference for "theoretical" terms and terms for unobservable entities pose quite *general* threats to any non-trivial form of SR. The overall worry can be summed up in slogan form. When it comes to the unobservable entities and properties posited in modern scientific theories, *What if it's all just empty talk?* Language-based considerations may lead us to think this is a serious worry, even today.

Let's start with meaning holism, championed by Quine and Davidson. The idea is that the meaning of any term in language, whether it be 'sofa', 'blue', 'oxygen' or 'electron', is not specifiable in some tidy definitional way, but rather is at least in part a function of the entire web of beliefs and practices that are accepted by the pertinent culture using those words. Therefore, when the beliefs and practices connected to a term change significantly - as they do, for example, across a scientific revolution that substantially changes the theory, world-view, and experimental practices in a science - the meaning of a word (say, 'electron', after the quantum revolution) is quite different than it was in the earlier epoch.[10] Kuhn famously argued that the meaning of 'mass' changes so substantially in the shift from Newtonian to Einsteinian paradigms that the respective uses before/after the relativity revolution literally do not refer to the same property.

If we are not careful, this view of meaning may lead us to think that if uses of 'electron' actually referred to anything in one epoch (either pre- or post-revolution), then they perforce cannot do so in the other. As Putnam described the worry in the 1970s:

> Let us suppose they are right, and that 'electron' in Bohr's theory (the Bohr-Rutherford theory of the early 1900s) does not refer to what we *now* call electrons. Then it doesn't refer to *anything* we recognize in present theory, and, moreover, it doesn't refer to *anything from the standpoint of* present theory (speaking from that standpoint, the only things Bohr *could* have been referring to were electrons, and if he wasn't referring to electrons he wasn't referring to anything). So if we use present theory to answer the question 'was Bohr referring when he used the term "electron" ?', the answer has to be 'no', according to Kuhn and Feyerabend. (1978, pp. 22-23, italics in the original)

In an obvious extension of the worry, one might further think: major theory changes keep happening in science, especially in physics where further revolutions are expected.

10 Minor changes or adjustments in our web of beliefs, of course, happen all the time, and presumably not even Kuhn and Feyerabend would regard them as really changing the meanings of the connected words.

So it is quite reasonable to suspect that future physicists will look back on *our* uses of 'electron' and judge that they were empty talk, referring to nothing. And since those future physicists will be in a much better epistemic situation than us, we should be worried indeed, already, that our assertions about electrons today are mere empty prattle; which is to say, anti-realists are right.

Now, nothing in the logic of the argument above depended on picking a term from physics, so in principle the threat applies to 'virus' and 'human Y-chromosome' just as much as 'electron'. Two things are worth remarking in this regard.

(1) One of the core claims of TSR is that we have reason to suspect that there will *not* in the future be any revolutions in biology that substantially change the meanings of 'virus' or 'human Y-chromosome' (and, indeed, that we can't conceive in any clear sense how such a revolution could occur). This already partly defuses the Kuhn-Feyerabend-holist worry sketched above.

(2) Since TSR is careful to exclude fundamental physics from its domain, one might think that we need not worry about physics revolutions changing the meaning of 'electron'. But, recall, TSR keeps at least certain basic, now-deeply-entrenched bits of physics lore outside of the quarantine zone; and that includes the lore that tells us that electrons exist, have a certain quantity of negative charge, flow through wires in electric circuits and from cathode to anode in cathode ray tubes, and so on. So we need to say more in order to protect our electrons from the present worry.

Holism is not a very popular view these days. But the core feature of holism that generates apparent problems for scientific realism, namely that the meaning (and hence reference, if any) of theoretical terms is given by associated descriptive content, is still defended by many, especially when it comes to terms for unobservable entities introduced in scientific theories. In other words, any version of *descriptivism* about meaning can lead to essentially the same threat to SR.

Chakravartty sketches the concern thus:

> Even in cases where successive theories ostensibly refer to the same kind of object (the electron, for example), changes in how it is described are sometimes so great that it may not seem entirely credible to maintain that each of these theories takes the same object as its subject matter. … Is it reasonable to assert that *o* exists, and yet be open to the possibility that one's conception of *o*, in terms of the set of properties a theory associates with it, may change, perhaps greatly? (2007, pp. 63-64)

For descriptivists, the meaning of a natural kind term or property term is typically given by a set of descriptions that are supposedly true of the denoted objects; or rather, if the term denotes anything in the world at all, it denotes whatever kind, entity or property is uniquely picked out by virtue of satisfying all, or some importance-weighted majority, of the associated descriptions. It is easy to see that this raises a serious problem if, over time, the scientific community makes major changes in the descriptions it associates with a term.

Descriptivism was for many years an almost universally accepted view of the meaning of any referential terms, singular or general; but the view was severely compromised by the arguments presented by Kripke in 1970, especially what have come to be known as the *ignorance* and *error* arguments. First, Kripke noted that speakers often associate to terms only descriptions that are insufficient to pick out a unique referent. Focusing on proper names, Kripke famously argued that most speakers associate with 'Feynman' something like 'famous physicist who worked at Caltech', a description which applied to Gell-Mann also. Similarly, it is likely that many users of the word 'platinum' associated that term with 'heavy, silver-colored, valuable metallic element', which fails to distinguish platinum from several other elements. Second, speakers may also associate descriptions that pick out the wrong individual or kind, e.g. associating 'first European to set foot in the New World' with 'Columbus', or 'light element which is a gas at normal room temperatures' with 'lithium', when lithium actually is a solid.

A variant of descriptivism has seemed, to some philosophers, able to evade Kripke's ignorance and error arguments: *cluster* descriptivism, first proposed in Searle (1958). According to cluster descriptivism, in a linguistic community a term is associated, not with a single description, but with a cluster of descriptions or attributes, and the reference of a use of the term is the individual or kind that uniquely satisfies a sufficient, weighted number of the properties in the cluster. The cluster need not be in each speaker's head as long as the community as a whole has a rich enough cluster, which seems to resolve cases of ignorance; and the cluster may contain some false descriptions as long as there are sufficient correct ones and they are among those that have greatest weight, which seems to resolve cases of error. But, as we argue in (2019), both ignorance and error arguments continue to afflict cluster descriptivism, even if the examples are less ubiquitous.[11] On the one hand the view falls prey to the ignorance argument. For example, when samples of platinum were first named by Europeans, any plausible cluster associated with the term would have applied also to at least palladium, and perhaps other metals as well. But the people who introduced and started to use the term were holding samples of platinum (with some impurities) in their hands and intending to refer to that substance and *only* that substance, and we think they succeeded. On the other hand, there is no guarantee that the cluster of beliefs associated with a term at any given moment in time will contain a core of weighty, correct beliefs about the extension of the term. The cluster of important characteristics initially ascribed to *Megalosaurus* by the early dinosaur researchers included these: amphibious, quadrupedal, and up to 22 meters long. Later decades revealed that Megalosaurs were land animals, walked upright on their hind legs, and grew to only 6 meters or so in length (see Cadbury 2001). Whatever descriptive content remained that was correct, in the cluster we might ascribe to the dinosaur research community, was surely not enough to entail referential success according to cluster descriptivism.

That there are changes in the cluster of beliefs we hold about kinds of things is obvious. That makes cluster descriptivism attractive, for cluster descriptivism seems to provide the required amount of flexibility that allows us to account for the preservation of referential continuity in spite of the change in our conceptions of things. The problem is that cluster descriptivism does not account adequately for continuity of reference and the problem arises when we confuse those flexible, clustery beliefs with reference-determining meanings. For, if the clusters change so much, and if the weights of the beliefs are shifted substantially, how are we to say, for instance, that previous members of the community were talking about the same things and kinds of things we do, and that they were wrong about the properties they attributed to them? The problems that the Kuhn-Feyerabend view raises for any form of realism come back in full force.

Granted, there are cases in which things work the way descriptivism or cluster descriptivism predict, cases where we realize that our conceptions were so muddled that we conclude there was nothing we were definitely referring to. The case of 'phlogiston' readily comes to mind. Some cases do work that way, but some others do not. Descriptivist theories maintain that reference is only possible via the mediation of a definite description or a cluster of descriptions that select the referent, and if nothing

[11] In Hoefer & Martí (2019) we argue that ignorance and error arguments continue to afflict more recent and sophisticated versions of cluster descriptivism, such as that advocated in Häggqvist & Wikforss (2018).

satisfies the description or the properly weighted cluster, reference fails. That is the way reference is *always* determined, according to descriptivists. But that is *not* how things always work, so we need an account that opens the door to the possibility that reference can be determined independently of the satisfaction of descriptions associated with terms.

It might be argued though that our confidence in the continuity of reference in spite of cases of substantial error or of ignorance is based on the fact that the cases that come to mind (platinum, *Megalosaurus*) are typically kinds of things with observable samples or that have left observable traces that are crucial in grounding reference. The case of 'electron', obviously, is rather different. But we will argue that descriptivism or cluster descriptivism does not give us the right account of reference in that kind of case either, so an alternative non-descriptivist account is required in the case of reference to, at least some, non-observable entities or phenomena. In the rest of this section we will examine the problems of descriptivist reference for unobservables such as electrons, and in the next section we explore how reference to the unobservables that TSR urges us to believe in does get established.

In the SR literature we keep coming back to the case of the electron because it illustrates so clearly the concern that can be raised about referential continuity across multiple, important theory changes. When J.J. Thomson talked about electrons in discussing his famous experiments, while accepting his Nobel prize in 1906, was he really talking about *the same things* as a modern physicist who mentions electrons while explaining the basic fermions in a graduate course on the Standard Model? We say yes, but under descriptivism it may not be possible to say this.[12] Ignorance and error problems threaten descriptivist reference to electrons in various ways.

Thomson and his peers in 1906 might have described the electron as a particle that composes cathode rays, has a certain constant mass and a certain constant negative charge (the exact values of which were not yet known), and contributes to making up atoms. So far, so good (by late 20th-century lights), we see here four bits of description that are still correct by modern lights.

But if we imagine pressing Thomson and his peers with further questions, things become murkier. We might ask "Do electrons have definite positions at all moments of time, and

[12] Under Kuhnian holism it is *clearly not* so. Multiple very important changes in belief about the electron, and about the nature of fundamental particles in general - as well as about mass, space, time, energy and more - occurred in physics between 1906 and, say, 1976, by which time the Standard Model was clearly established. Some descriptivists, anti-realist descriptivists, would of course be happy to side with Kuhn on this point.

follow continuous trajectories through space?", to which they would surely have answered "Yes, of course.", whereas a contemporary physicist would answer "Definitely not!" We might go on and ask Thomson and his peers: "Does an electron have a wave-like nature? Does it have intrinsic angular momentum? Might it have further internal structure or sub-components?" They would probably have answered these questions (not without a certain level of irritation and bewilderment) with *No, No, Maybe but I doubt it.* The 1976 physics lecturer, however, would answer these questions as follows: *Yes, Yes, Certainly not.* This is a substantial amount of disagreement, about things physicists of both eras would agree are extremely important, fundamental, properties that electrons either do or do not have! Here we see the error argument looming.

And there are other illustrations of how large the shadow of the error argument looms. Referential continuity across multiple revolutions is implicitly assumed for all the particles discovered in the late 19th century and first half of the 20th (protons, neutrons, positrons, various mesons, etc.) despite sometimes major shifts in beliefs about their nature. To mention just two: protons went from being assumed fundamental to being thought to be composed of three quarks; and throughout a substantial portion the 20th century, neutrinos were supposed to be massless. That property was at the very core of the characterization of neutrinos. It was quite a surprise to find out in 1998 that neutrinos had, after all, mass. But no one thinks (we hope) that neutrinos do not exist, or that we were not referring to neutrinos before 1998, merely on the basis of the fact that, as far as we can tell, nothing satisfies the description neutrinos were long thought to satisfy, nor even just the set of the weightiest, (allegedly) definitory properties associated with 'neutrino'.[13]

Going back to Thomson and electrons, a cluster descriptivist might argue that referential continuity exists between Thomson's day and ours, if she maintains that the *core, central, most important* bits of description attached to the term 'electron' are: "composes cathode rays, has a certain constant mass and a certain constant negative charge, and contributes to making up atoms". The last three, at least, would be reckoned as indeed central, and the middle two (constant mass and charge, as per the now-known values) even essential to being an electron, by physicists today.

13 Neutrinos may belong to the quarantine zone of fundamental physics rather than the safe/established part, and if so one is free to doubt their existence for the same reasons that one doubts other elements of fundamental theory. Our point is that we hope no one doubts the existence of neutrinos because they considered, before 1998, 'Neutrinos have no mass' to be true by definition. We thank Ana Maria Cretu for drawing our attention to the example of neutrinos. For discussion of potential implications of this case for essentialism, see Cretu (2018).

We have two responses to this. First, the mere fact that we can, with hindsight, pick out a "core" set of descriptive associations for 'electron' that were not false and that were sufficient to uniquely pick out electrons, will not do. If it is to be a genuine theory of reference, descriptivism must allow the "core" of highest-weighted descriptions to be specified prospectively, not with 70 years' hindsight; in other words, the descriptions in the cluster have to be the descriptions that the users of the term, *at the time*, do associate with the term, and consider to be "core" or "weightiest". It is by no means clear that Thomson and his peers would have nominated each of these four descriptions of electrons to be part of the core, or that they would not have added further descriptions now known to be false.

Second, it is not hard to slightly fictionalize the early history of the electron so as to make the situation impossible to salvage under descriptivism. Imagine that Thomson and other physicists had had the view (based on perhaps a too-direct reading of experimental results?) that the mass and charge of electrons are variable within a certain range, rather than uniformly all the same. And imagine that the consensus had been, for some years, what certain prominent physicists *did* think at first: that electrons are *not* components of atoms. But otherwise, let's stipulate, most of the experimental knowledge is the same as it was in the late-1800s. Then the only non-false (by late 20th century lights) descriptions in the core/cluster of beliefs held by Thomson and his peers in 1906 might have been: negatively charged particle that comprises cathode rays. And now the ignorance argument bites, because in the cathode ray tubes of the day, there were also various negatively charged ions traveling from cathode to anode or screen (Broadway & Pearce (1939)). Even ignoring the erroneous parts of the cluster, the true component left would not pick out electrons uniquely.

Any form of descriptivism will have trouble arguing that fictional-Thomson and his peers referred to anything at all when they talked about 'electrons', if we today are more or less right in our beliefs about what *we* call 'electrons'. But we think that fictional-Thomson and his peers would still have been talking about electrons, just as we are today, despite the many false beliefs we would say they had concerning them. The intuition that this is right is not substantially weaker, we think, in the fictional history scenario than it is in the real scenario described earlier.

There is no denying that in the very early history of the use of 'electron' there may have been, and in fact was, referential indeterminacy, and any adequate theory of reference needs to make room for that. In fact, and as we will argue in the next section, accepting some degree of indeterminacy is crucial in order to understand the complexity of the establishment of a referential practice. But by 1906 Thomson and his peers were definitely referring to electrons, insufficient or erroneous characterizations notwithstanding.

Finally, it is worth remarking that, just as it is easy to craft a situation in which the majority of descriptions associated with 'electron' in the early years after the introduction of the term are false, it is also easy to imagine - at a superficial level, at any rate - future revolutions that render false a majority of the descriptions *we now* associate with 'electron'. How could this be? Well, we may imagine that a future theory attributes internal structure to electrons after all, and also postulates a slow change, over cosmic time, in the mass/charge ratio of the electron and most other massive particles. Worse, future theory might tell us that electrons actually come in two subspecies with distinct internal structural features that lead to *slightly* distinct, but very similar, observable properties. (One of the two subspecies might be vanishingly rare in our solar system, explaining why past experiments never clearly revealed the existence of both sub-kinds.) And so forth.

To be clear, we do not claim that these future twists in the story of the electron are *actually* genuinely conceivable (without adverting to radical skeptical scenarios), i.e., that a real possible-physics with these features, empirically at least as adequate as current physics, exists in conceptual space. We merely claim that, given the actual history of physics in the past two centuries, it would be rash to claim that this can be ruled out with certainty. This, of course, is why - according to TSR - fundamental physics *theory* belongs in the quarantine zone. What is *not* any longer in the quarantine zone according to TSR is the fact of the *existence* of electrons, and the rough or approximate truth of the *basic lore* concerning them; we have too many epistemic handles on these facts to be able to conceive their falsehood without invoking radical skeptical scenarios.[14]

In sum, descriptivism plays into the hands of scientific anti-realism by raising the prospect that Thomson was perhaps not talking about anything real at all when he wrote his famous papers on electrons, and – worse – that the same may still be true for us when *we* talk about electrons. Cluster descriptivism may seem to provide a way to elude the danger, but in fact it does not. But if descriptivism is not the right way to understand how terms for unobservable entities refer, what is? How does the reference of a word like 'electron' get established?

[14] As we will argue below, we should not give in to the temptation of thinking that those epistemic handles form a kind of descriptivist cluster. As mentioned above, epistemic handles are not simply bits of descriptive knowledge; this will be clarified further in section 4.

4. Reference to theoretical natural kind and property terms

Neither classical nor cluster-descriptivism gives us a correct account of the reference of terms, including theoretical terms. The alternative to descriptivism is the so-called causal-historical approach, originally introduced by Kripke (1980), Donnellan (1970) and Putnam (1973), and later developed by Michael Devitt in his (1981) and other works. The causal-historical approach works well for proper names and for natural kind terms denoting observable, easily ostendible medium-sized goods. Moreover, it has seemed to be attractive also as an account of reference to unobservable entities and magnitudes, because while such things cannot be easily ostended, it has always been part of the causal-historical approach that terms can be introduced via definite descriptions (as in the case of 'Neptune' discussed by Kripke). And once a term is introduced, whether by baptism, ostension of paradigmatic samples, or by description, the capacity to refer is passed on, and maintained through the subsequent chain of users of the term, so that even if theories later change (as happened in the case of electrons), it is easy to maintain that users of the term are still talking about the same things. In some cases, the introduction of a term may involve both description and ostension, e.g. Benjamin Franklin might have introduced 'electricity' by the description: "Whatever physical kind or process underlies *these* phenomena" (ostending lightning and the subsequent discharge phenomena induced in his famous kite experiments).

But many philosophers of science have worried that the causal-historical approach goes too far in the direction *away* from descriptivism, making reference too easy and causing different problems with prominent cases in the history of science. 'Phlogiston' is usually invoked in this regard. Common sense dictates that 'phlogiston' never referred to anything at all; but if it had been introduced by means of the description "substance crucially involved in combustion phenomena", then arguably it would have referred to *oxygen* – making subsequent decades of scientific debate into complete nonsense (unbeknownst to the participants in the debate and the experimental work). This example is typically taken to show that one cannot avoid introducing theoretical descriptions of a more detailed and robust kind into the story of how reference to unobservable theoretical entities and magnitudes is established (see for example Kroon (1985), Psillos (1999, 2012)).

In this section we will defend the causal-historical approach to the reference of terms for theoretical entities. We will use as a foil the causal descriptivism defended in Psillos

(2012), a hybrid account that combines aspects of descriptivism and aspects of the causal-historical picture. The reason to focus on Psillos' approach is that it is, in principle, especially promising: on the one hand, the incorporation of some descriptivist tenets is meant to free the theory of reference from the excessive referential success that is supposed to plague the causal-historical approach and, on the other hand, the incorporation of aspects of the causal-historical picture seems to allow the view to evade the problems that affect descriptivism.[15] However, as we will argue, Psillos' approach is still vulnerable to the kinds of problems that cluster descriptivism faces. We will argue that TSR meshes perfectly with the causal-historical approach and resolves the concerns that have been raised about its ability to work well for theoretical terms for unobservable entities and magnitudes.

According to Psillos, two conditions must be satisfied for a theoretical term *t* to refer to an entity or magnitude *x*: (a) *x* must be the cause of a phenomenon φ (intuitively, the phenomenon scientists wish to explain with the help of *t*); (b) *x* satisfies a description D(*x*) that captures the ways in which the posited referent is supposed to be causally connected to the phenomenon φ.[16] When *t* is introduced, reference is not successfully established unless both conditions are satisfied:

> If causal description D(*x*) is satisfied by some entity *x*, but *x* is not the cause of φ, then *t* does not refer to it. Conversely, if something does indeed cause φ, but D(*x*) is not satisfied by this something, then *t* does not refer to it.[17] (p. 222)

The causal-historical approach to reference has two distinct parts: one of them deals with the introduction of the term, the other one deals with its subsequent uses. Psillos' proposed referential mechanism is clearly meant to apply to the introduction of a term *t*. It is not entirely clear to us whether Psillos means for the two conditions to apply in subsequent uses, down the chain of communication so to speak, so that, whenever the

15 A number of philosophers have argued that a causal component to the account of reference is needed to avoid problems of referential failure or problems of referential continuity; see, for example, Newton-Smith (1981) chapter 7.
16 Are (a) and (b) explicitly held in the mind of the introducer of the term *t*? Or are they in the minds of the relevant community (whether explicitly or implicitly)? We will assume the latter, since it seems to be the most plausible way of interpreting the proposal. We believe that nothing in our discussion will depend on this question.
17 Notice that *causation* appears in both of Psillos' conditions, and this seems to be his reason for describing his view as a form of "causal descriptivism". Psillos' account is nonetheless quite different from some earlier accounts of reference that have gone by the same moniker, for example in Lewis (1984) and Jackson (1998), but is somewhat closer to Kroon (1987). We will not discuss the earlier causal descriptivism; for criticism, see Raatikainen (2006) and Martí (2020).

term *t* is uttered, it will refer to *x* only if the two conditions are satisfied, no matter how much the beliefs that the community associates with *t* have changed. Some of the things Psillos says seem to suggest that this is so and, in fact, continuity of reference would seem to require that both conditions, i.e., *x* causing φ and *x* satisfying D, be in place. However, some of Psillos' remarks suggest that although condition (a) definitely has to be correct and retained, there can be refinements and changes in D(*x*) over time, and that it may contain some erroneous parts at first.[18] Psillos recalls David Papineau's elegant way of putting the distinction between core parts of D(*x*) which presumably *must* be retained (the 'Yes' parts), those that may or may not be retained ('Perhaps'), and those descriptions that definitely do not belong in D(*x*) ('No'). We claim that, as long as there is a core part of D(*x*) that must be correct on pain of referential failure, the view is vulnerable to the typical problems descriptivist views face.

Recall our fictional-Thomson from section 3. *All* of his central beliefs about electrons were incorrect, except for the belief that that they were negatively charged and present in cathode rays. Even if we take the relevant phenomenon φ associated with 'electron' to have been cathode rays, we noted that there were other negatively charged particles in the rays at that time, negative ions of various atoms and compounds. So this may be seen as a case of referential indeterminacy or reference failure, and Psillos' causal descriptivism does not deliver the desired referential continuity in this case. Admittedly, this is a fictional case, but other problematic cases are not hard to find in real history of science. Consider again the case of neutrinos. At the time of its introduction, 'neutrino' seems to fail to meet both of Psillos' conditions, (a) and (b). Regarding (a), neutrinos were not introduced as the *cause* of any phenomenon φ; they were introduced to "balance the books" of conservation principles, being the bearers of the energy and angular momentum that seemed to be missing in beta decay processes. To say that neutrinos *cause* energy conservation, or cause it to not be violated, would be an abuse of language.[19] But there was no other known, observable phenomenon of which neutrinos were thought to be the cause, to play the role of φ in (a). Regarding (b), neutrinos were initially

18 See his discussion of the history of chlorine, pp. 227-228.

19 In so far as physicists think of anything as being a *cause* of energy conservation, it would be the time-translation symmetry of the fundamental Lagrangians of particle physics, or perhaps the time-translation symmetry of spacetime's structure.

A similar issue regarding condition (a) arises for the introduction of 'Higgs boson': it is not introduced as the *cause* of fact that some elementary particles have mass; the word usually used is 'responsible'. Even if this can be understood as causation in some extremely broad sense, it clearly is not the kind of *causal contact* between the named entity and the namer that is meant to help ground reference according to causal descriptivists.

thought to have either zero or very low rest mass; but theoretical developments in the 1940s and experimental results in the 1950s seemed to show that neutrinos must have zero rest mass (Murayama 2002). By the late 1950s, this property would have been put into the 'Yes' part of any plausible D(x) capturing the community's beliefs, so this looks like a potential case in which the "refinement" of D(x) leads to referential failure.[20]

Restoring to descriptions a crucial role in the determination of reference solves the specific problem that 'phlogiston' raises, but it risks creating potential new problems with historical (and contemporary) cases. Moreover, it sacrifices one of the key insights of the causal-historical approach: that once reference is up and running, we can coherently imagine discovering that *any* description that scientists had in mind at the time of introducing a term turns out to be incorrect. But, how does reference get up and running? We contend that the causal-historical approach has all the resources required to handle reference to theoretical terms for unobservable entities and magnitudes.

First of all, we need to set aside a couple of misconceptions concerning the causal-historical approach that have made it appear less flexible than it actually is. One concerns the sense in which causation is involved. The causation involved concerns reference transmission, i.e., how reference in later uses depends on the existence of a causal chain of reference-borrowing that leads back to the period of uses that fix or establish reference. There is no requirement that the entity or magnitude named by a new term "be the cause" of the introduction, in any sense, nor that it be the cause of some phenomenon, the desire to explain which prompts the introduction – although sometimes one or both of these things will be the case. And this means, secondly, that as Raatikainen (2007) stresses, the causal-historical approach is perfectly compatible with a term being introduced with only a definite description initially offered to fix the term's reference, as in the famous example of Neptune.[21]

But if a theoretical term is introduced purely by means of a description, doesn't this mean that on the causal-historical approach, the term forever more designates whatever entity satisfies the description (if any), and hence that the description cannot fail to apply? No,

[20] Psillos' remarks about the changes in D(x) over time make this concern particularly acute, because he thinks that "Ultimately, the reference should be fixed by the kind-constitutive properties …" (p. 228), and having rest mass (or not) is surely a kind-constitutive property for fundamental particles.
[21] It is true that Michael Devitt (1981) offers an account in which both the introduction and the transmission are causal. But as he himself points out (2015) the account is meant to apply to paradigm proper names and natural kind terms that have ostendable samples. And even in those cases, Devitt agrees, the names in question can be introduced via a definite description, a description that nevertheless does not play any role in the transmission of the capacity to refer.

because the *establishment* of reference, getting a referential practice up and running in a linguistic community, is in most cases a *process* that takes place over a period of time, not something accomplished at the very moment of introduction of a term. As Martí (2015) puts it,

> … we should not think that bestowing a name is an act; it is a process. It requires success in launching a practice, and launching practices is not something that occurs instantly. … The fact that the idealized picture of naming presented by Donnellan and Kripke focuses on the very simple cases where there is a dubbing ceremony should not distract us from the variety of cases in which the process of introducing a name in a language is considerably more complex … (p. 87)

If this is true for the case of proper names, as we believe it is, imagine how much more scope for complexity there is in the establishment of reference for theoretical terms for unobservable entities and magnitudes!

Many things can happen after a term is first introduced, whether the introduction involved only descriptive beliefs, or rather some element of ostension ("Let's call the things making *those* trails in our cloud chamber '*t*'."), or some combination of both (and perhaps more). Once we abandon the idea that introducing a name (be it of a person or an unobservable) consists in a single well-defined *act* of naming, and we embrace the view that reference-establishment is a process, it is much easier to account for that enormous complexity. It is our view that the causal-historical approach should, and can easily, incorporate the idea of naming as a process and thus be able to handle paradigmatic cases of referential failure, success, and continuity in the history of science.[22]

If a theoretical term is introduced with only a description, then at that point reference may be *so far* not established (if the description picks out nothing real – e.g., 'phlogiston' and 'substance given off by things that burn in the process of combustion'); or it may be tentatively established (e.g., 'Neptune', 'large planet beyond Uranus whose gravitational effects on Uranus explain the discrepancies in its orbit given Newtonian gravity'), but this is not the end of the story. Even starting from such an impoverished position, scientists invariably have further ideas about how to find out about the named thing – how to produce it, for example, or block it, or how to determine its mass; what other effects it

[22] That the introduction and establishment of a theoretical term is an extended process and also a complicated one, with many twists and turns before clarity and stability are eventually achieved, will be no news to historians of science who study the history of how 'oxygen', 'electron' and many other scientific terms came to have an established use.

may produce; and so forth. As the community goes about its business, theorizing and experimenting and observing, some of these ideas may lead to the community achieving new *epistemic handles* on the entity. When things go well, those epistemic handles will both solidify the community's belief in the existence of the entity, and potentially also change the community's beliefs about the entity's nature and properties. It is the achievement of consensus about the epistemic handles we have on an entity that makes for an established referential practice. But by the time that is achieved, or at some later time, the first descriptions associated with the term may be dropped from the community's beliefs. For example, perhaps some day physicists will discover, and announce, that "The Higgs boson turns out to have nothing to do with the origins of mass!" – even though the Higgs was introduced to explain why W and Z gauge bosons have mass, much as 'Neptune' was introduced to name the theorized planet that explained anomalies in the orbit of Uranus. This is perfectly imaginable if in the future the production of Higgs bosons in particle accelerators becomes routine, and our manipulations of them become entangled with other phenomena and other practices; those epistemic handles might come to be much more important to physicists' referential practices involving 'Higgs boson' than the initial theoretical description assigning the Higgs boson (or field) responsibility for the fact that certain particles have rest mass.[23]

Epistemic handles play two crucial roles in our approach. On the one hand, they are fundamental to our defense of TSR: when they are numerous and diverse enough, they give the scientific community grounds to be certain of the truth of some belief or the existence of some entity. On the other hand, they play a role in the way in which a referential practice becomes consolidated in a community. This role is not necessarily one that guarantees the existence of the entity, or the truth of the community's beliefs about it! Rather, the epistemic handles that contribute to established referential practices ensure that the community has *some* agreed forms of evidence for the existence of the entity or magnitude named by the term, and agreed ways for gathering further evidence for it or information about it. These "ways of knowing" may include things like: microscopes to visually image it; detectors to detect its presence or measure its properties; "factories" to produce more of it; and indirect effects that its existence or presence should

[23] Note that in this scenario, our current theoretical description of the Higgs turns out to have been wrong, in line with TSR's insistence that fundamental physics theory is still in quarantine. It might be correct to say, in this scenario, that the Higgs *particle* is no longer in the quarantine zone, as we have asserted above regarding electrons – we do not have to take a stand on this question. Even if it *would* be correct to regard the Higgs particle as out of quarantine, this would not imply that the (updated) theoretical beliefs about it also have that status.

induce, where we have ways of observing or knowing whether these effects occur. And clearly, this list is not exhaustive.[24]

But even when a community has enough handles on a theoretical entity or magnitude to clearly meet the threshold for having an established referential practice, it may not yet have enough handles, with enough interconnection to other parts of the community's knowledge, to make the entity or magnitude's existence indubitable in the sense of TSR.

Before a referential practice is established for a theoretical term, our approach makes room for both indeterminacy of reference and referential failure, even for a term that, like 'oxygen' or 'neutrino', will eventually be successful. Referential indeterminacy is a real phenomenon and it has to be acknowledged and accounted for by any theory of reference. Indeterminacy may arise in more than one way. There are cases in which it may be entirely unclear, even with hindsight, whether a term referred in the early stages of its use and, if so, to what. The word 'oxygen', introduced by Lavoisier, may illustrate this phenomenon, since at the beginning Lavoisier believed that he was dubbing a substance that was the principle of generation of acids, which is false, and of course he had no idea that the oxygen generated as the end product of certain chemical processes was a diatomic molecule (O_2) rather than a type of elementary atom, or even a continuous fluid substance.

TSR acknowledges the existence of temporary referential indeterminacy and, in our view, so does the causal-historical picture, even though the simple examples that its proponents have often used, with their insistence on dubbing ceremonies, obscure that fact. Some aspects of the 'Madagascar' case, often discussed in the literature on proper names, exhibit the same kind of temporary referential indeterminacy. Since no official dubbing of the island occurred (and there was a previous time when 'Madagascar' was used, but not as the name of the island), it is likely that there was some period of time for which we ought to say that it is not entirely clear what a particular utterance of the name referred to. Nowadays things are settled and it is entirely clear what 'Madagascar' refers to, but a grey area, a period of indeterminacy, needs to be accepted. Michael Devitt (1981, 2015) has

24 As we have noted earlier, epistemic handles are not simply bits of descriptive knowledge or beliefs (see footnotes 7 and 14). A proponent of cluster descriptivism might argue that the handles can be described, and hence those descriptions can be part of a descriptive cluster. The answer to this is that *anything* can be described. The interactions of the parts of an automobile engine can be described, but the descriptions are not what make the car move. Epistemic handles can be described, but it is the handles themselves, not descriptions of them, that play a role in established referential practices (and in our knowledge more generally).

explained the 'Madagascar' case from the causal-historical point of view in terms of multiple grounding and partial reference. Even mistakes can ground reference anew, providing a new reference for an old name. The consolidation of the grounding is a process, and processes hardly ever have clear dividing lines.

Let's now come back to the case that was supposed to be difficult for the causal-historical approach: phlogiston. The word was introduced with an attempted initial fix on its reference via a description, something like 'substance whose presence in certain materials explains their combustibility, and which is given off by those materials during combustion'. This description does not pick out anything, of course. Certain experiments and alleged connections of phlogiston with other phenomena (such as metallic appearance) led to the creation of some putative epistemic handles on phlogiston.[25] But these handles unraveled as further experimental knowledge increased, which is why an established referential practice was never achieved and the community eventually settled on the view that phlogiston does not exist. This historical case is extremely interesting, but we do not see anything in it that challenges either TSR or the causal-historical approach to reference for theoretical terms, as we understand it.[26]

Of course since the success of reference to an unobservable depends, on our view, on many factors that are external to the minds of users, the approach would seem to face a serious problem which can be put in the form of a question: how does one identify what one is talking about, and even whether one is talking about anything at all? The point of the approach is precisely that 'one' does not identify it. If a term that has been introduced has connected successfully to something, and the term is being passed to other links in the chain of communication that use it, as the approach holds, with the intention to refer

25 These putative epistemic handles correspond to the reasons often cited by scientific anti-realists for the claim that phlogiston theory constitutes a historical example of successful prediction and explanation without truth and hence a problematic case for defenders of IBE – that is to say, for most proponents of other versions of SR, but not for the proponent of TSR.
26 All this does not exclude the possibility of a fictional scenario something like this: Lavoisier is not around or at any rate does not coin the term 'oxygen'; as tensions grow in the experiments on phlogiston somebody makes the bold proposal that phlogiston is actually not *given off* in combustion, but *acquired,* and that it is present in normal air; that phlogiston has nothing to do with the shiny appearance of metals but instead affects their appearance when metals get "phlogisticated" (oxidized); and so on. In such a scenario a referential practice for 'phlogiston' might have been successfully established, in which the word refers to what we now call *oxygen*. Chang (2012) argues that had 'phlogiston' been retained, it might have come to refer to chemical potential energy, or perhaps to free electrons. That referential practices can stabilize, after a period of initial indeterminacy, in more than one way, is consistent both with the causal-historical approach to reference and with TSR.

to the same thing, the uses of the term refer to that thing. That success is not dependent on what is in anyone's head.

The descriptivist may insist: this is a deeply unsatisfactory view about reference: you do not know if you are referring nor, if you do, to what. It all depends on whether someone was lucky and their original introduction or subsequent uses managed to connect term and referent. So, successful reference depends on some form of referential luck. But observe that whenever an external condition is constitutive of the success of a process or an action, the agent's point of view is not sufficient to gauge the success. For instance, the justified-true-belief analysis of knowledge introduces an external element to the epistemic success of knowing vs. merely believing: truth, an element that in general cannot be determined just by looking into one's mental states. The situation as regards reference is no different.

5. Implications for perspectival scientific realism

In recent years a number of philosophers of science have advocated views that can be grouped under the umbrella term 'perspectivism', and some of these (e.g., Giere (2006), Teller (2020) and Massimi (2018)) advocate views that merit the more specific label 'perspectival scientific realism' (PSR). As is almost always the case, each philosopher has their own specific set of views that are not quite the same as those of any other philosopher, and this makes characterizing the varieties PSR difficult. We will not attempt this task, but rather mention some characteristic claims that one can find advocated by one or more proponent of PSR, in order to get a feel for the view, before making some closing remarks about how TSR compares to PSR, particularly in light of the considerations about reference developed in section 4.

Perspectivism can be seen as opposed to a "view from nowhere" understanding of scientific knowledge. In a variety of ways, scientific knowledge is to be understood as perspective-bound or perspective-relative. In particular, it is to be understood as thoroughly *historical* (that is, it emerges in specific human communities and from specific historical - hence, contingent - goals, concerns, concepts and practices). Nevertheless, the advocate of PSR maintains that science certainly gives us truths, but always *from within a perspective.* But once one has mentioned the perspectival qualification, one can go on to say that some bits of scientific lore, even about unobservable entities, are *really true.* For example, from the perspective characteristic of 20th century Western chemistry, it is simply true to say that diamonds are made of carbon atoms bound together in a certain type of crystalline configuration - full stop. But we are not to take this as a *universal* truth that somehow transcends any or all perspectives.

We find many of the things that perspectivists say to be clearly right and also unproblematic from the broader standpoint of defending a non-trivial form of scientific realism. We acknowledge that the meaning of scientific terms gets shaped through historical processes, by people with certain aims and certain existing ideas and practices at their disposal. We also acknowledge a point stressed by Teller (2020): scientific realists advocate belief in the truth-*or-approximate-truth* of certain parts of scientific lore, but the very notion (vague) of 'approximate' truth can only be understood relative to certain goals and interests, i.e., from a perspective.

That said, arguably, perspectives bear some family resemblance to Kuhnian paradigms, and it may be tempting to read the claim that scientific statements are perspective-bound as similar, if not identical, to Kuhn's claim of the incommensurability of the meanings of (similar- or identical-looking) statements made within very different paradigms.[27] If that is the intention, then we wish to point out that the inadequacy of descriptivism (and the closely related Kuhn-Feyerabend holism we discussed in section 3) as accounts of meaning and reference suggests limitations on the extent to which it is right to see scientific statements as perspective-bound.

If one starts from a holist or a descriptivist notion of what determines the reference of natural kind and property terms used in science, it is natural to think that when two communities with quite different perspectives and beliefs both assert a sentence like 'Electrons have rest mass m_e', despite superficial appearances they are neither saying the same thing, nor necessarily either agreeing or disagreeing with each other. One community may associate the key terms 'electron', 'mass' with very different descriptions than the other, and connect these terms with different complexes of theoretical belief. But if the causal-historical account is, as we believe, the right way to understand the reference of many terms in scientific discourse, then our scientific language (and also the language of everyday life) possesses certain "pitons" that anchor our discourse firmly to reality outside of our heads.

As we made clear in section 2, TSR claims that the *existence* of electrons is now beyond doubt (barring radical skepticism), and let us suppose that the same can be said about the

27 This similarity is noted by many writers on perspectivism, and no doubt explains why advocates take pains to point out that perspectivism is not simply a form of truth relativism.

fact that there is a natural property/magnitude that we call 'rest mass'.[28] Then, even though it remains possible that we will in the future dramatically change some of our theoretical beliefs about electrons, and about rest mass, nevertheless the statement '*Electrons exist and have rest mass* m_e' is no longer hostage to our theoretical beliefs, nor in any important sense perspective-bound. It is instead an objective fact about our world, in as strong a sense as any scientific realist should hope for.

To say this is not to deny that there may be many statements - perhaps even a majority - in science whose meanings are *in part* fixed by holism-style connections to other parts of accepted lore and by the practices and interests of the relevant community. But it is important also to acknowledge that the causal-historical approach to reference imposes a certain amount of externalism, giving us a partial grip at least on reality that is independent of our current concepts and perspectives.

The considerations about meaning and reference we raised in sections 3 and 4 above, together with the central claims of TSR, militate against some of the more radical interpretations of perspectivist theses, those that are close to a Kuhnian/constructivist view; but they are, we believe, perfectly compatible with many moderate perspectivist theses. We believe that perspectivism can coexist perfectly well with the causal-historical approach to reference, and even with TSR, but it will be tempered with a somewhat greater amount of trans-perspectival truth than its adherents might currently expect.

Acknowledgments

Versions of this paper were presented at the Dubrovnik Philosophy of Science Conference (2015) and the Language and Realism Conference (Barcelona 2017). We thank the audiences for their comments. We are also grateful to Ana Maria Cretu, Michela Massimi, Stathis Psillos and Panu Raatikainen. We are also very grateful to two anonymous referees for this journal, whose comments prompted what we think are important improvements in the paper.

Funding

The research for this paper was supported by projects FFI2016-76799-P and FFI-2015-70707-P of the Spanish MEIC.

[28] We will assume, as seems plausible, that the meaning of mathematical expressions (such as the number represented by m_e) and other basic lexical items of English such as 'has' or 'is' have remained constant.

References

Broadway, L.F. and Pearce, A.F. (1939): 'Emission of negative ions from oxide cathodes'. *Proceedings of the Physical Society* 51 (2): 335-348.

Cadbury, D. (2001): *The dinosaur hunters*. London. Fourth Estate.

Callender, C. (2019): 'Can we quarantine the quantum blight?' in *Scientific Realism and the Quantum*, S. French & J. Saatsi (eds.), Oxford University Press: 57-77.

Chakravartty, A. (2007): *A metaphysics for scientific realism*. Cambridge University Press.

Chang, H. (2012): *Is water H_2O? Evidence, realism and pluralism*. Dordrecht. Springer.

Cretu, Ana-Maria (2018): *What good is realism about natural kinds?* PhD Thesis, University of Edinburgh.

Devitt, M. (1981): *Designation*. Columbia University Press.

Devitt, M. (2015): 'Should proper names still seem so problematic?'. A. Bianchi (ed.): *On Reference*. Oxford University Press: 108-143.

Giere, R. (2006): *Scientific perspectivism*. University of Chicago Press.

Häggqvist, S., & Wikforss, Å. (2018). 'Natural kinds and natural kind terms: myth and reality'. *The British Journal for the Philosophy of Science* 69 (4): 911-933. http://dx.doi.org/10.1093/bjps/axw041 (accessed 20-May-2020).

Hoefer, C. (2019): 'Scientific realism without the quantum' in *Scientific Realism and the Quantum*, S. French & J. Saatsi (eds.), Oxford University Press: 19-34.

Hoefer, C. & Martí, G. (2019): 'Water has a microstructural essence after all'. *European Journal for the Philosophy of Science* 9 (12). https://doi.org/10.1007/s13194-018-0236-2 (accessed 20-May-2020).

Kripke, S. (1980): *Naming and necessity*. Harvard University Press.

Kroon, F.W. (1985): 'Theoretical terms and the causal view of reference', *Australasian Journal of Philosophy* 63 (2): 143-166.

Martí, G. (2015): 'Reference without cognition'. A. Bianchi (ed.): *On Reference*. Oxford University Press: 77-92.

Martí, G. (2020): 'Names, descriptions and causal descriptions. Is the magic gone?'. *Topoi* 39 (2):357-365.

Massimi, M. (2018): 'Four kinds of perspectival truth'. *Philosophy and Phenomenological Research*, 96 (2): 342-359.

Murayama, H. (2002): 'The origin of neutrino mass'. *Physics World*, 15 (5): 35-39.

Newton-Smith, W.H. (1981): *The rationality of science*. Routledge Press.

Psillos, S. (1999): *Scientific realism.* Routledge Press.

Psillos, S. (2012): 'Causal descriptivism and the reference of theoretical terms'. A. Raftopoulous and P. Machamer: *Perception, realism and the problem of reference*. Cambridge University Press: 212-238.

Putnam, H. (1973): 'Meaning and reference'. *The Journal of Philosophy*, 70 (19): 699-711.

Putnam, H. (1978/2010): *Meaning and the moral sciences.* Routledge Press.

Raatikainen, P. (2006): 'Against causal descriptivism'. *Mind & Society* 5, 78-84.

Raatikainen, P. (2007): 'Theories of reference and the philosophy of science'. Manuscript (accessed 20-May-2020): https://philpapers.org/rec/RAATOR

Searle, J. (1958): 'Proper names'. *Mind* 67: 166-173.

Stanford, K. (2006): *Exceeding our grasp: science, history, and the problem of unconceived alternatives*, New York: Oxford University Press.

Teller, P. (2020): 'What is perspectivism and does it count as realism?' in *Understanding perspectivism: scientific challenges and methodological prospects,* M. Massimi & C. McCoy (eds), Routledge Press, 2020.

Tulodziecki, D. (2017): 'Against selective realism'. *Philosophy of Science*, 84 (5): 996-1007.