How and How Not to Be Whiggish About ‘Phlogiston’

Jonathon Hricko

Abstract: Understanding the semantics of theoretical terms from past science involves determining what, if anything, they referred to. Some ways of assigning referents to such terms are Whiggish, in the sense that they introduce anachronisms that distort the past, while others are not. My aim in this paper is to develop a non-Whiggish semantic theory, one that avoids Whiggish reference assignments. In order to do so, I make use of the example of ‘phlogiston.’ I argue that it would be Whiggish to maintain that ‘phlogiston’ sometimes referred to free electrons, but not Whiggish to maintain that it sometimes referred to hydrogen. And I argue that we can avoid Whiggish reference assignments by focusing on the operations that scientists perform. On the theory I defend, theoretical terms refer by means of the operations scientists use to identity their putative referents.

Keywords: Reference, Theoretical Terms, Whig History, Anachronism, Operations, Phlogiston

1 Introduction

For some time, philosophers have been interested in the semantics of theoretical terms, and the challenges of referring to unobservable entities. Many philosophers have taken an interest in this issue because of its implications for the scientific realism debate. If our best theories are approximately true, as scientific realists claim, then their theoretical terms must refer to unobservable entities. A related, but distinct, reason to take an interest in this issue
is that it has implications for our understanding of the history of science, and this will be my focus in this paper.

One of our primary ways of learning about the history of science is by reading old scientific texts. And these texts often contain a number of out-of-date theoretical terms (e.g., ‘phlogiston,’ ‘caloric,’ and ‘luminiferous ether’). It can be challenging to understand these texts, and semantics can play an important role in helping us to understand them. For this reason, Thomas Kuhn (1982a, p. 674) claims that “historians dealing with old scientific texts can and must use modern language to identify referents of out-of-date terms.” And Philip Kitcher (1993, p. 100) maintains that “using the ideas of modern science to provide a picture of the relationships between [past scientists] and the world” is not “Whiggish, illegitimate, [or] question begging,” on the grounds that, “without some picture of the relationship between past figures and the world, history is impossible: we cannot frame any hypotheses about what their discourse means.” These philosophers, who have themselves shaped the scientific realism debate in many important ways, thus recognize a role for semantics outside of the debate—it can help us to understand past science.

With this goal in mind, I’ll focus on retrospective reference assignments to the theoretical terms of past science. In section 2, I’ll take some cues from some recent and not-so-recent literature on the historiography of science, and I’ll distinguish Whiggish reference assignments, which distort the past in various anachronistic ways, from non-Whiggish reference assignments, which do not. I’ll use this distinction to draw another one, namely, between Whiggish and non-Whiggish semantic theories. In section 3, I’ll consider a historical example, and argue that it would be Whiggish to maintain that ‘phlogiston’ sometimes referred to free electrons, but not Whiggish to maintain that it sometimes referred to hydrogen. In section 4, I’ll argue that the theory of reference that Stathis Psillos (1999) defends is an example of a Whiggish semantic theory. Contrary to what Psillos claims, his theory entails the Whiggish result that ‘phlogiston’ referred to free electrons. Finally, in section 5, I’ll propose my own non-Whiggish semantic theory, which makes use of Hasok Chang’s (2009a; 2011) notion of operational meaning. According to the theory I’ll defend, theoretical terms refer by means of the operations scientists use to identity their putative referents.
2 Whig History and Whiggish Semantics

My first goal is to explain what Whiggish semantics is. Since I’m borrowing the term ‘Whiggish’ from the historiographical literature on Whig history, I’ll start there. I’ll discuss what Whig history is, where it came from, and why historians of science think it is so objectionable. After doing so, I’ll show that some of the historiographical issues surrounding Whig history are also semantic issues that concern retrospective reference assignments to theoretical terms of past science. In that case, it also makes sense to talk about whether reference assignments, and the semantic theories that yield them, are Whiggish.

2.1 Whig History

The term ‘Whig history’ has its source in Herbert Butterfield’s *The Whig Interpretation of History*. Butterfield characterizes Whig history as involving “[t]he study of the past with one eye, so to speak, upon the present,” and he goes on to claim that it is “the source of all sins and sophistries in history, starting with the simplest of them, the anachronism” (1965/1931, pp. 31–32). Butterfield chose the term ‘Whig’ to characterize this particular interpretation of history because the histories told from the perspective of the English Whig party exemplify the kind of history he wanted to attack. Whig historians, in the literal sense, viewed history as a progressive process that culminated in the beliefs and ideals of the Whig party. And they distinguished between those historical actors who anticipated or contributed to that progress, and those who hindered it. Butterfield generalizes from this literal usage, and uses the term ‘Whig’ to refer to any kind of history that views the past as a kind of inevitable, progressive march toward the present.

It took some time for Butterfield’s terminology to catch on among historians of science. But as Nick Jardine (2003, p. 127) observes, by the mid-1970s, it became common for them to apply the terms ‘Whig’ and ‘Whiggish’ to works that depict past science in anachronistic ways. Butterfield (1965/1931, p. 24) himself was concerned primarily with a kind of anachronistic value judgment by which historians select what was important in the past by appealing to what we consider important today. Obviously, there is no guarantee that our value judgments will line up with those of the past actors we study. And so, this kind of anachronism of selection may lead us to emphasize the wrong aspects of some historical episode, and to thereby
misrepresent it. As Jardine (2004, p. 261) notes, historians of science have followed Butterfield, and they use the term ‘Whiggish’ to label this kind of anachronism of selection.

But Jardine also points out that historians of science have used the same term to label another kind of anachronism that is conspicuous by its absence from Butterfield’s discussion, namely, conceptual anachronism. Jardine (2003, p. 127) characterizes this kind of anachronism as “the application of our categories to the works and deeds of those who lacked such categories.” And he identifies it with the kind of anachronism of which Quentin Skinner (1969) was so critical (Jardine, 2004, p. 261). Skinner argued that such anachronism can lead historians to “understand the agent to be doing something which he would not—or even could not—himself have accepted as an account of what he was doing” (1969, p. 6; quoted in Jardine 2004, p. 261). Historians of science have condemned conceptual anachronism, on the grounds that ‘understanding’ an agent in this way amounts to misunderstanding that agent. As Jardine notes, historians of science have generally sought to combat conceptual anachronism by replacing “Whiggish and anachronistic narratives” with “historically sensitive” ones, i.e., ones that adhere to “actors’ explicit categories” (2004, p. 268).

Jardine himself adopts a more measured view of anachronism. He maintains that history sometimes requires conceptual anachronism in order to facilitate “communication of the resultant insights to living readers,” and “to analyse and explain as well as describe past deeds and works” (2004, p. 262). Moreover, he argues that historians can often depart from their strict adherence to actors’ categories without committing vicious anachronism, which he understands as “historically incoherent interpretation of past deeds and works” (2000, p. 252). In order to distinguish cases in which the retrospective application of a category results in vicious anachronism from

---

1. To take one example, this is presumably what Adrian Wilson and T. G. Ashplant (1988, p. 11) have in mind when they argue that “[t]he case of whig history demonstrates that it is perfectly possible for a historian to work with inappropriate categories, with anachronistic assumptions and expectations, in short with a present-centred mental ‘set’, and so to produce history which distorts the past.”

2. Jardine is not alone in this regard. Other recent examples include Nick Tosh (2003) and Chang (2009b). Butterfield himself is a not-so-recent example, since, as Chang (2009b, p. 255) notes, he turned to Whig history later in life. And even Kuhn (1977/1968, p. 110), when suggesting that “the historian should set aside the science that he knows,” claimed that historians should do so “[i]nssofar as possible,” and that “it is never entirely so, nor could history be written if it were.”
cases in which it does not, Jardine (2000) appeals to presuppositions. He
observes that there are presuppositions involved in the application of a cate-
gory. And he argues that if those presuppositions were not realized at some
time in the past, then it would be viciously anachronistic to apply that cate-
gory to actions taking place at that time. However, if the presuppositions are
realized, then although it may be anachronistic to apply a category that past
actors lacked, it is not viciously anachronistic. In fact, Jardine maintains
that departing from actors’ categories can even enhance our understanding
of the past in some cases.

Some examples should make the idea clear. To take one example, Jar-
dine claims to “have only the mildest reservations” about talk of Galileo’s
‘socio-professional identity’ (2000, p. 251). Even if those in Galileo’s time
lacked that category, the presuppositions for its application were arguably
realized, and so a retrospective application of that category does not result in
a vicious anachronism. In contrast, for Jardine, talk of Aristotle’s ‘biological
investigations’ is viciously anachronistic (2000, pp. 261–262). Since biology
did not exist as a discipline until some time in the nineteenth century, the
presuppositions required for the application of the category did not exist in
Aristotle’s time. Jardine is clear that the issue is not a linguistic one con-
cerning the use of terms like ‘biology’ and its cognates. Indeed, he maintains
that his view differs from others “in not requiring that disciplinary terms be
applied only to the activities of agents who themselves apply those terms to
their activities” (2000, p. 262).

My primary concern is not with the retrospective application of categories
to past works and deeds, but with a related kind of anachronism, to which
I’ll turn momentarily. Like Jardine, I take it that a complete prohibition
on anachronism would be unwise, and so I’ll attempt to distinguish between
vicious and non-vicious cases of that kind of anachronism. And given that the
kind of anachronism with which I am concerned bears some resemblance to
the kind with which Jardine is concerned, I’ll indicate the similarities between
Jardine’s way of distinguishing vicious from non-vicious anachronism and my
own.

2.2 Whiggish Semantics

The kind of anachronism with which I am concerned has to do with retro-
spective reference assignments to theoretical terms of the past, examples of
which include ‘phlogiston,’ ‘caloric,’ and ‘luminiferous ether.’ Understand-
ing scientific discourse involving such terms requires us to assign referents to those terms (where reference failure is a possible reference assignment). And oftentimes it will conduce to our understanding to specify the referents of those terms in modern language—after all, it won’t be particularly illuminating to be told that ‘phlogiston’ refers, if at all, to phlogiston. This use of modern language obviously opens the door to anachronism, and just as in the cases discussed above, such anachronistic reference assignments can be vicious or non-vicious. My goal at this point is to distinguish reference assignments that are viciously anachronistic from those that are not, and to use this distinction in order to distinguish between Whiggish and non-Whiggish semantic theories.

It will be helpful to discuss an example first. Kitcher (1993, p. 100) has argued for a particular retrospective reference assignment, namely, that Joseph Priestley’s ‘dephlogisticated air’ referred, at least on some occasions, to oxygen. However, Christina McLeish (2005, p. 680) objects to this reference assignment, on the grounds that “providing a semantics for statements made in the past is not the same as writing a history of that past, even though the history is very important to the semantics.” History, according to McLeish, aims to tell us what happened. Providing a semantics for statements made in the past involves determining whether past actors correctly described what happened, or referred to it at all. She admits that our present-day conceptual schemes can play a valid role in determining what happened in the past. In that case, we can say truly that Priestley isolated relatively pure samples of oxygen gas. When it comes to semantics, McLeish holds that the question of whether past actors referred to what happened is a question, not just about what happened, but about the representational states of those past actors. She makes the further point that it would be a mistake to let our present-day conceptual schemes stand in for the representational states of those past actors. In that case, the fact that Priestley isolated relatively pure samples of oxygen gas is not sufficient to establish that he referred to oxygen when he used the term ‘dephlogisticated air.’ And while our present-day conceptual schemes might be sufficient for us to refer to oxygen, if they were unavailable to Priestley, then Kitcher must find some other way to make it plausible that Priestley’s ‘dephlogisticated air’ referred to oxygen. McLeish goes on to conclude that this particular reference assignment is implausible.

In section 5, I’ll argue that it’s not viciously anachronistic to hold that Priestley’s ‘dephlogisticated air’ referred, at least on some occasions, to oxygen. That said, even if I think McLeish is wrong about this particular ref-
erence assignment, her objection to Kitcher’s view gives us the resources for distinguishing reference assignments that are viciously anachronistic from those that are not. In particular, she makes two points that I will employ in order to draw this distinction.

First of all, there is McLeish’s point that reference is not simply a matter of what happened in a speaker’s immediate environment, but depends on the representational states of the speaker. The insight here is that it takes some work for scientists to fix the reference of theoretical terms to theoretical entities, especially to unobservable ones. It is, of course, well known that ostension is often sufficient to fix the reference of a term in ordinary contexts. But in scientific contexts, it’s often more difficult to fix reference. John Worrall (2007, p. 148) brings out the difficulty with the following example:

Think about how one might ‘ostend’ the electromagnetic field, say, in order to ‘baptise’ it: clearly we could only know that we are ostending the field (in fact you can point in any direction you like!) via the *theory* of the field.

I take Worrall’s point that something more than ostension is required to fix the reference of many theoretical terms. But it’s important to recognize that, in addition to theory, experiment can also play a role in fixing reference, as I’ll argue in section 5. For now, the point I want to make is that, even if fixing reference in ordinary contexts doesn’t require much sophistication, fixing the reference of theoretical terms requires scientists to have some rather sophisticated beliefs and/or abilities. Their representational states therefore depend, at least in part, on various theoretical beliefs and/or experimental capabilities.

Secondly, there is McLeish’s point that it would be a mistake to substitute the representational states of past speakers with our present-day conceptual schemes, at least if we have good reason to think that past speakers lacked those conceptual schemes. While I take it that McLeish is correct, the point that she makes applies to more than just conceptual schemes. More generally, it would be a mistake to attribute our theoretical beliefs and experimental capabilities to past scientists who lacked them.

At this point, we have what we need to distinguish reference assignments that are viciously anachronistic from those that are not. Consider a scientist $S$ who uses some theoretical term $t$. In order for $S$ to use $t$ to refer to some theoretical entity $x$, $S$ needs to have some rather sophisticated beliefs and/or abilities. Now let’s suppose that we conclude that $S$ used $t$ to refer
to $x$. If drawing this conclusion requires attributing to $S$ beliefs and/or abilities that $S$ actually lacked, then the assignment of $x$ as the referent of $t$ is a *viciously anachronistic reference assignment*. In what follows, I’ll also refer to such a reference assignment as a *Whiggish reference assignment*. Otherwise, we have a *non-Whiggish reference assignment*, and these come in two varieties. If we specify the referent of $t$ in modern language that was unavailable to $S$ (say, by using our present-day theoretical terms), then we have an *anachronistic-but-non-vicious reference assignment*. And if we specify the referent using language available to $S$, then we have a *non-anachronistic reference assignment*. Finally, a *Whiggish semantic theory* is one that entails at least one Whiggish reference assignment, while a *non-Whiggish semantic theory* entails none.

My way of distinguishing reference assignments that are viciously anachronistic from those that are not owes something to Jardine’s way of distinguishing between vicious and non-vicious anachronism in the application of categories to the works and deeds of past actors who lacked those categories. I’m in agreement with Jardine that characterizing the past in terms of modern language unavailable to past actors is not always viciously anachronistic, and can often enhance our understanding of the past. And I take it that my use of beliefs and abilities of past scientists plays much the same role as Jardine’s use of presuppositions. While it may be possible to subsume my distinction under his, it would, at the very least, take some work, since Jardine does not address the reference of theoretical terms. In any case, the important point, for my purposes, is not that my distinction ultimately differs from Jardine’s, but that we have a way of distinguishing reference assignments that are viciously anachronistic from those that are not.

Before moving on, it will be helpful to make some clarifications. First of all, it’s worth emphasizing that I’m treating reference failure as a possible reference assignment. While it’s difficult to conceive of how reference failure would ever be Whiggish, it may be the case that, for some theoretical terms of the past, the only non-Whiggish reference assignment is reference failure.

Secondly, our candidates for referents of theoretical terms, both past and present, are entities that we presently take to exist. But that doesn’t mean that all retrospective reference assignments (or at least the non-empty ones) are anachronistic. After all, it’s possible that some past scientists shared at least some of our present-day terms, beliefs, and abilities, in light of which they succeeded in referring to entities that we presently take to exist.

Thirdly, it’s worth noting that there may be a point to Whiggish referen-
ence assignments. For example, even if it were Whiggish to maintain that Priestley’s ‘dephlogisticated air’ referred to oxygen, it may be useful to put forward this reference assignment for pedagogical purposes. It may be a good first approximation or useful fiction that students can dispense with later in the semester, in favor of a more sophisticated view. However, even if they’re useful for some purposes, Whiggish reference assignments are still historiographical mistakes.

Fourthly, and finally, determining whether a particular reference assignment is Whiggish is a non-trivial task. It will involve investigating the language, beliefs, and abilities of past scientists. It will also involve determining what combination of language, beliefs, and abilities is required for fixing reference to a particular theoretical entity. As such, it will often be a matter of some debate whether a particular reference assignment is Whiggish. That said, I take it that these are among the issues that historians and philosophers of science ought to debate.

3 How to be Whiggish About Phlogiston

My goal, in the remainder of the paper, is to develop a non-Whiggish semantic theory, i.e., one that yields only non-Whiggish reference assignments. In order to meet this goal, I’ll make use of one example of a Whiggish reference assignment, and one example of a non-Whiggish reference assignment, which I introduce in this section. Both are reference assignments to the term ‘phlogiston.’ To a first approximation, phlogiston was a substance that chemists working in the seventeenth and eighteenth centuries posited as a component of metals and combustible substances. Metals were thought to lose phlogiston when they transformed into what chemists at the time called ‘calxes,’ which we now know as oxides; and combustible substances were thought to lose phlogiston in the process of combustion. I’ll now argue that it’s not Whiggish to maintain that ‘phlogiston’ sometimes referred to hydrogen, but that it is Whiggish to maintain that it sometimes referred to free electrons.

3.1 ‘Phlogiston’ and Hydrogen

The first reference assignment I’ll consider is that ‘phlogiston’ sometimes referred to hydrogen. A number of philosophers have considered this proposal, which is one subject of an exchange between Philip Kitcher, Mary Hesse, and

Hesse (1982, p. 707) responds as follows:

We have not only to say that phlogiston sometimes referred to hydrogen and sometimes to absorption of oxygen, but we have to convey the whole ontology of phlogiston in order to make plausible why it was taken to be a single natural kind.

Kuhn (1982b, p. 712) agrees with Hesse, and provides the following response:

The processes to which she refers are independent, and the older literature of the history of science provides countless examples of the ease with which one may complete the first [i.e., assigning referents to ‘phlogiston’] without taking even a step toward the second [i.e., explaining why phlogiston was taken to be a single natural kind]. The result is an essential ingredient of Whig history.

In what follows, I’ll ignore the points about the absorption of oxygen, and about the second process, and I’ll simply consider whether the proposal that ‘phlogiston’ sometimes referred to hydrogen is a Whiggish reference assignment in the sense I proposed in section 2.2.

In order to do so, I’ll first have to discuss the reasons for taking this proposal seriously. In short, it’s a reasonable proposal to consider because a number of phlogiston theorists identified phlogiston with inflammable air, which we now call hydrogen. Richard Kirwan is arguably the chemist who did the most to defend this identification, and he argues for it in the following way. He begins by listing a number of phlogiston’s properties:

By phlogiston is generally understood that principle in combustible bodies on which their inflammability principally depends; that principle to which metals owe their malleability and splendor; that which combined with vitriolic acid [i.e., sulfuric acid] forms sulphur; that which diminishes respirable air. (1782, p. 197)

---

3The subject of the exchange here is really Kitcher’s earlier assertion that a number of phlogiston theorists “believed that the inflammable air which they had isolated was phlogiston,” and “went on to record the properties of inflammable air (hydrogen) using the term ‘phlogiston’” (1978, p. 534). I’ll discuss some of these phlogiston theorists in detail below.
In the pages that follow this passage, Kirwan goes on to argue that these properties are all properties of inflammable air. He concludes that “inflammable air . . . is one and the same substance with phlogiston” (1782, p. 209). Some years later, in his *Essay on Phlogiston and the Constitution of Acids*, Kirwan is clear that this identification is not an idiosyncrasy of his phlogiston theory, but that it has “met the approbation of the most distinguished philosophers, both at home and abroad” (1789, p. 5). He goes on to list a number of phlogiston theorists who, he claims, also accept this identification, including “Dr. Priestley, Mr. Bewly, Mr. Bergman, Mr. Morveau, De La Metherie, Chaptal, Crell, Wiegleb, Westrumb, Hermstadt, Kaersten, &c.” (1789, p. 5). To be sure, not all phlogiston theorists maintained this identification; and among those who did, some subsequently abandoned it. That said, the fact that many prominent phlogiston theorists maintained the identification is surely enough to motivate us to take seriously the proposal that ‘phlogiston’ sometimes referred to hydrogen.

Now that I’ve discussed the reasons in favor of this proposal, we can ask whether it’s Whiggish. My basic idea is that, because Kirwan and others identified phlogiston with inflammable air, which just is hydrogen, this reference assignment is not Whiggish. That said, there are two reasons why one might suspect that it really is Whiggish. First of all, I haven’t shown that it’s not Whiggish to maintain that ‘inflammable air’ referred to hydrogen. Secondly, it may be the case that, even though some chemists at the time identified phlogiston with inflammable air, and others identified inflammable air with hydrogen, none identified phlogiston with hydrogen. With regard to the first concern, I’ll have more to say about why it’s not Whiggish to maintain that ‘inflammable air’ referred to hydrogen later in this section, and also in section 5. But this reference assignment would be a side issue if, in response to the second concern, one could show that there were some chemists at the time who used ‘hydrogen’ (or one of its cognates) to specify the referent of ‘phlogiston.’

I’ll now discuss two such cases. The first is implicit, and is found in Antoine Lavoisier’s response to the following passage from Kirwan’s *Essay*:

> it seems to me very difficult to conceive how pure air can unite to phlogiston . . . without forming a new compound . . .; it seems, therefore, more reasonable to conclude, either that it forms water, as Mr. Cavendish thinks, or fixed air, as I shall endeavour to prove in the following sections. (1789, p. 39)
In the commentary that he provided for the second edition of Kirwan’s *Essay*, Lavoisier writes: “[Kirwan] thinks that, from this combination of oxigene [sic] and hydrogene [sic], there does not always result water; that in some circumstances the result is fixed air, or carbonic acid” (in Kirwan, 1789, p. 56). Lavoisier could not have characterized Kirwan’s view in this way unless he identified Kirwan’s pure air with his own oxigene, and Kirwan’s phlogiston with his own hydrogen. The second case is more explicit, and occurs about twenty years later. In the course of speculating on the nature of metals, Humphry Davy writes of “the adherence of their phlogiston or hydrogene [sic]” (1808, p. 364), thus explicitly identifying the two. Hence, at least some chemists at the time identified phlogiston and hydrogen, in which case it’s not Whiggish to maintain that ‘phlogiston’ sometimes referred to hydrogen.

At this point, one might question whether it’s Whiggish to assign hydrogen as the referent of Lavoisier’s and/or Davy’s ‘hydrogene.’ After all there’s a strong sense in which Lavoisier’s hydrogen is different from our hydrogen, since Lavoisier’s hydrogen gas was a compound of hydrogen base and caloric (Lavoisier, 1965/1789, p. 198). But, to echo a point that Kuhn (1996/1962, p. 55) makes about oxygen, it’s surely too much to conclude that chemists could not refer to hydrogen until sometime after the 1860’s, when the caloric theory of heat was finally abandoned. To anticipate a point that I’ll make in section 5, this is because chemists working in the late eighteenth century were already able to isolate relatively pure samples of hydrogen gas. Moreover, they were able to determine enough of the properties of those samples in order to conclude that they were dealing with a substance that differed from other gases with which they were familiar.

In that case it’s not anachronistic, and therefore not Whiggish, to maintain that ‘phlogiston’ (along with ‘hydrogene’ and ‘inflammable air’) sometimes referred to hydrogen. At worst, this reference assignment involves an anachronistic application of modern language (‘hydrogen’) to specify the referent of ‘phlogiston,’ but does not involve attributing beliefs and abilities to past chemists who lacked those beliefs and abilities. And at best, this reference is not even anachronistic, since past chemists did specify the referent of ‘phlogiston’ by using a cognate of ‘hydrogen’ that is still used today, namely, the French term ‘hydrogène.’
3.2 ‘Phlogiston’ and Free Electrons

The second reference assignment that I’ll consider is inspired by Hasok Chang’s work on the Chemical Revolution. While Kuhn’s remarks suggest that Whig history is rather easy to fall into, Chang (2009b; 2012, ch. 1) argues that it’s a more difficult attainment, and that historians have not been properly Whiggish about phlogiston. He attempts to show what a truly Whiggish view of phlogiston would look like, not as an idle exercise in Whiggishness, but in order to highlight the unexplored potential of the phlogiston theory. One of the conclusions that he draws is that, “[i]f we were to be truly whiggish, we would recognize phlogiston as the precursor of free electrons” (2012, p. 44). The second proposal that I’ll consider, then, is that ‘phlogiston’ sometimes referred to free electrons.

I’ll motivate this proposal by appealing to the two ways in which Chang supports the identification of phlogiston with free electrons. First of all, Chang compares phlogistonist explanations of various phenomena with those explanations that we currently accept. Regarding the common properties of metals, Chang (2012, pp. 43–44) writes:

Phlogistonists explained the common properties of metals by saying that all metals were rich in phlogiston; ... the phlogistonist account actually has a close resonance with the modern notion that all metals share metallic properties because they all have a “sea” of free electrons.

Regarding the processes of reduction and oxidation, which phlogiston theorists understood as phlogistication and dephlogistication, respectively, Chang (2012, p. 44) quotes the following passage from the chemist Gilbert Newton Lewis (1926, pp. 167–168):

If they [the phlogistonists] had only thought to say “The substance burning gives up its phlogiston to, and then combines with, the oxygen of the air,” the phlogiston theory would never have fallen into disrepute. Indeed, it is curious now to note that not only their new classification but even their mechanism was essentially correct. It is only in the last few years that we have realized that every process that we call reduction or oxidation is the gain or loss of an almost imponderable substance, which we do not call phlogiston but electrons.
And regarding “the production of flame in combustion,” Chang (2012, p. 45) notes that “flame is a plasma, which is mostly a mixture of positive ions and electrons.” He argues that the identification of phlogiston with electrons “fits in nicely here, if we take the release of flame as a result of the dissociation of phlogiston (electrons) from the combustible substance” (2012, p. 45). Hence, when it comes to these phenomena, the phlogistonist explanations are similar to the explanations that we currently accept.

The second way in which Chang supports his identification of phlogiston with free electrons involves showing that a number of phlogiston theorists posited a close connection between phlogiston and electricity. Among their reasons for doing so was that they could transform calmes into metals by means of electricity, and they understood such transformations in terms of gain in phlogiston (2012, p. 44). Chang mentions that there were at least 23 theorists who posited such a connection. I’ll briefly note three examples. First of all, Chang (2012, p. 44) mentions John Elliott (1780, p. 92), who proposed to replace the term ‘phlogiston’ with ‘electron’ because of the close connection between phlogiston and electricity. Secondly, Chang (2012, pp. 80–82) discusses Joseph Priestley (1802, p. 202), who, after recounting some experimental results, claims that they “favour the hypothesis of two electric fluids, the positive containing the principle of oxigen [sic], and the negative that of phlogiston.” And thirdly, Chang (2012, p. 80) notes that George Smith Gibbes (1809, p. 13) posits a similar connection when he claims that “[t]he principle of the negative side of the galvanic apparatus resides in all combustible bodies, . . . and answers exactly to the Phlogiston of Scheele.” For these reasons, Chang (2012, p. 44) concludes that the connection between phlogiston and electricity is not just a “retrospective fabrication.”

Of course, even if it is not, Chang is correct about the Whiggishness of identifying phlogiston with free electrons. And although Chang does not put the point in terms of reference, it’s also the case that it is truly Whiggish to maintain that ‘phlogiston’ sometimes referred to free electrons. To be sure, we can admit that chemists, at the time, could use the term ‘phlogiston’ to refer to electricity, which we now understand in terms of electrons. But to say that these chemists could thereby refer to free electrons is to attribute to them a theoretical sophistication that they surely lacked, and that wasn’t attained until the late nineteenth century at the earliest. J. J. Thomson’s work on cathode rays, along with H. A. Lorentz and Pieter Zeeman’s work on the Zeeman effect, arguably constitute the earliest moments at which reference to electrons was fixed. Their experimental capabilities and theoretical ex-
planations were simply unavailable to phlogiston theorists working a century earlier. Insofar as reference to electrons requires something approximating the beliefs and abilities of physicists working in the late nineteenth century, it would be viciously anachronistic to maintain that chemists working in the late eighteenth century could refer to electrons. Hence, this second proposal amounts to a Whiggish reference assignment.

4 An Example of Whiggish Semantics

I'll now consider one semantic theory for assigning referents to theoretical terms, which is due to Stathis Psillos (1999, ch. 12). I'll show that, contrary to what Psillos himself argues, his theory can be used to support the claim that ‘phlogiston’ sometimes referred to free electrons. In that case, it yields a Whiggish reference assignment, and is an example of a Whiggish semantic theory.

Psillos aims to develop a causal-descriptivist theory of reference that captures the strengths, but not the weaknesses, of purely causal theories and purely descriptive theories, respectively. He begins with the familiar observation that descriptive theories make reference too difficult (1999, pp. 281–282). Such theories require that the referent of a theoretical term is that entity which satisfies all or most of the descriptions associated with that term. As is well known, Saul Kripke (1980) and Hilary Putnam (1975) showed how reference can succeed in the absence of such a description. This was taken to be a desirable result, since the descriptions that scientists associate with various theoretical terms are often flawed. To take a couple of examples: Lavoisier’s oxygen was supposed to be the principle of acidity, and John Dalton’s atoms were supposed to be surrounded by spheres of caloric. If we want to say that they nonetheless succeeded in referring, then reference to theoretical entities can’t merely be a matter of satisfying a description.

In response, a number of philosophers developed various causal theories of reference, and Psillos goes on to make another familiar observation about such theories, namely, that they make reference too easy (1999, p. 290). To a first approximation, these theories maintain that reference to a theoretical entity is fixed by first specifying some phenomena, and then introducing a theoretical term as a name for the entity that is the underlying cause of those phenomena. But so long as those phenomena have some cause or other, the theoretical term cannot fail to refer—it will refer to the cause, whatever it
happens to be. If we want to say that terms like ‘caloric’ and ‘luminiferous ether’ failed to refer, then reference can’t merely be a matter of specifying the cause of some phenomena.

While I’m generally in agreement with Psillos’s assessment of the respective weaknesses of descriptive and causal theories, I aim to argue that his own causal-descriptivist semantic theory is Whiggish. I’ll now turn to that theory, which Psillos (1999, pp. 295–296) characterizes in the following way:

a theoretical term \( t \) typically refers by means of a core causal description of a set of kind-constitutive properties, by virtue of which its referent \( x \) is supposed to play a given causal role in respect of a certain set of phenomena. Given this, the following conditions are easy to motivate.

1. A term \( t \) refers to an entity \( x \) if and only if \( x \) satisfies the core causal description associated with \( t \).
2. Two terms \( t' \) and \( t \) denote the same entity if and only if (a) their putative referents play the same causal role with respect to a network of phenomena; and (b) the core causal description of \( t' \) takes up the kind-constitutive properties of the core causal description associated with \( t \).

Psillos understands kind-constitutive properties as “those whose presence in an item makes that item belong to a kind,” and as “those whose presence makes a set of objects have the same, or sufficiently similar, manifest properties, causal behaviour and causal powers” (1999, p. 288). These are the same properties that allow the kind to fulfill its causal role in the production of a given set of phenomena (1999, pp. 297–298). Psillos also claims that, “[i]f the kind-constitutive properties attributed to the referent are not the causal origin of the information associated with the term employed to refer to it, then the term fails to refer” (1999, p. 291). The information associated with the term is the core causal description mentioned above. And in order for the kind-constitutive properties to be the causal origin of this information, it must be the case that those properties are, in fact, causally responsible for producing the phenomena in question.

Psillos relies on two examples of theoretical terms in order to illustrate his theory: ‘luminiferous ether,’ and, to a lesser extent, ‘phlogiston.’ Regarding the former, he argues that his theory yields the result that ‘luminiferous
ether’ referred to the electromagnetic field (1999, p. 296). It’s worth noting that this could be a Whiggish reference assignment, though I won’t argue for this claim here. Instead, I’ll focus on his other example, namely, ‘phlogiston,’ and I’ll argue that Psillos’s theory yields the Whiggish result that it referred, at least sometimes, to free electrons.

Psillos himself argues that his theory yields the result that ‘phlogiston’ fails to refer. His goal in doing so is to show that his theory amounts to an improvement over a purely causal theory of reference, which yields the counterintuitive result that ‘phlogiston’ refers to oxygen. It does so, according to Psillos, because “[p]hlogiston’ was introduced on many occasions by means of a causal description, i.e. one that singled out phlogiston as the physical magnitude causally involved (given off) in combustion” (1999, pp. 290–291). Psillos goes on to explain why such a theory yields this counterintuitive result:

If we follow the letter of the causal theory and accept that ‘phlogiston’ was coined to refer purely existentially to whatever is causally involved in combustion, then the conclusion is inescapable—‘phlogiston’ refers to oxygen, since the latter is what is causally involved in combustion. (1999, p. 291)

In contrast, Psillos maintains that “[w]hat it is correct to say is that ‘phlogiston’ refers to nothing” (1999, p. 291). When it comes to his own theory, he claims that he “can explain why ‘phlogiston’ does not refer,” on the grounds that, “if the core description fails, then the putative entity does not exist” (1999, p. 298). Psillos holds that the core causal description associated with ‘phlogiston’ will at least include a description of “the property that it is released during the process of combustion” (1999, p. 298). His all-things-considered view regarding ‘phlogiston’ is that this core causal description fails to pick out anything, and so ‘phlogiston’ fails to refer.

While Psillos may succeed in showing that, on his theory of reference, ‘phlogiston’ fails to refer to oxygen, he doesn’t succeed in showing that it fails to refer altogether. As we saw in section 3.2, there is something released during the process of combustion, namely, free electrons. In that case, there is something that has the kind-constitutive property that, according to Psillos, the core causal description associated with ‘phlogiston’ must describe. To be sure, the core causal description will describe other kind-constitutive properties of phlogiston, e.g., that it is a component of metals, and that it is released in the process of dephlogistication, i.e., oxidation. But as we saw in
section 3.2, these are also properties of free electrons. In that case, the core causal description of phlogiston’s kind-constitutive properties is satisfied by free electrons. Moreover, this also shows that, when it comes to phenomena like combustion and oxidation, phlogiston and free electrons play the same causal role. And so, on Psillos’s theory, ‘phlogiston’ actually refers to free electrons. In that case, it’s an example of a Whiggish semantic theory.

5 Towards a Non-Whiggish Semantics

In this final section, I’ll develop and defend a non-Whiggish semantic theory. My goal in this section is to make it plausible that we can avoid Whiggish reference assignments if we pay attention, not just to what scientific theories say, but also to what scientists do. In other words, we ought to pay more attention to scientific practices, and in particular, to the experimental operations that scientists perform. In order to develop my non-Whiggish alternative to Psillos’s theory of reference, I’ll make use of Chang’s (2009a; 2011) notion of operational meaning, which I’ll discuss momentarily. And after putting forward my alternative, I’ll show that my approach yields the non-Whiggish result that ‘phlogiston’ sometimes referred to hydrogen, and avoids the Whiggish result that it sometimes referred to free electrons.

To begin with, I’ll discuss Chang’s notion of operational meaning, which he illustrates in terms of the example of Lavoisier’s ‘oxygen’ (2011, pp. 415–420). For Lavoisier, oxygen gas is a compound of oxygen base and caloric, and oxygen base is his principle of acidity, i.e., that which makes acidic substances acidic. Both of these claims played a central role in Lavoisier’s oxygen theory, and both were abandoned by chemists sometime in the nineteenth century. Nonetheless, chemists retained oxygen, and Chang is concerned with the question of why they were justified in doing so. His answer is that, while the theoretical meaning of ‘oxygen’ changed over time, there was continuity at the operational level. He explains the basic idea as follows:

All of the procedures that Lavoisier had used for producing and identifying oxygen gas are still repeatable and valid; . . . Heat some red oxide of mercury intensely; collect the evolving gas in a glass jar; see things burn with special vigor in that gas, and animals live longer; breathe it and feel a lightness in the lungs; explode it together with hydrogen gas and make water. This operational
stability is what is responsible for fixing the extension or reference of “oxygen”, to the extent that it has been fixed over the centuries. (2011, p. 419)

Chang goes on to use this idea in order to show how distinct theoretical terms from distinct theories can share an operational meaning. He claims that “it is a straightforward matter to observe that the operational meaning of Lavoisier’s ‘oxygen’ was pretty much the same as that of Joseph Priestley’s ‘dephlogisticated air’, or Carl Wilhelm Scheele’s ‘fire air’” (2011, p. 419). All three chemists disagreed with one another over various theoretical issues, but they shared many of the same procedures for producing oxygen gas. And if the reference of Lavoisier’s ‘oxygen’ was fixed by such procedures, then in light of these shared procedures, Priestley’s ‘dephlogisticated air’ and Scheele’s ‘fire air’ referred to oxygen as well.

Incidentally, this conclusion provides some support for a couple of claims I made in previous sections. First of all, in section 2.2, I indicated my preference for Kitcher’s view over McLeish’s when it comes to the reference of Priestley’s ‘dephlogisticated air.’ If the reference of Lavoisier’s ‘oxygen’ was fixed by means of experimental operations that he shared with Priestley, then surely Priestley’s ‘dephlogisticated air’ referred to oxygen by virtue of those very same operations. Secondly, in section 3.1, I indicated my view that both Kirwan’s ‘inflammable air’ and Lavoisier’s ‘hydrogene’ referred to hydrogen. The same story that Chang tells about Lavoisier’s ‘oxygen’ applies equally well to Lavoisier’s ‘hydrogene.’ And if Lavoisier’s ‘hydrogene’ referred to hydrogen in light of various experimental operations that he shared with Kirwan, then surely Kirwan’s ‘inflammable air’ also referred to hydrogen.

While Chang doesn’t present his notion of operational meaning as a way to avoid Whiggish reference assignments, I take it that a focus on operations holds the key for doing so. If operations are the means of fixing the reference of theoretical terms, and if these operations were shared by past scientists, then we don’t have to attribute to past scientists any beliefs or abilities that they lacked in order to conclude that they succeeded in referring. And in that case, none of the reference assignments that I just discussed is Whiggish.

I’ll now put forward my non-Whiggish alternative to Psillos’s theory of reference, which I’ll present in a form similar to the form that Psillos uses. I propose that a theoretical term $t$ refers by means of the operations that scientists use to identify the putative referent of $t$. And given that, I’ll motivate the following two conditions:
1 A term $t$ refers to an entity $x$ if and only if scientists can use the set of operations associated with $t$ to identify $x$.

2 Two terms $t'$ and $t$ denote the same entity $x$ if and only if (a) there are operations associated with $t$ that are also associated with $t'$; and (b) scientists can use those operations to identify $x$.

According to condition 1, if there is no set of operations for identifying the putative referent of a theoretical term, then reference fails. This may be the case if the putative referent is hypothetical, in the sense that scientists aren’t sure whether it exists, or if the putative referent has been eliminated from science because scientists have shown that it doesn’t exist. Condition 2 is meant to show how terms from distinct theories, or even distinct historical periods, can refer to the same entity.

Condition 2 is really the condition that is supposed to ensure that reference assignments to past terms are not Whiggish. The basic idea is that we can avoid Whiggish reference assignments by first examining the operations that past scientists associated with their theoretical terms, and then looking at present science to determine what, if anything, these operations are sufficient for identifying. If they are sufficient for identifying something in the current ontology of science, then it’s possible to make retrospective reference assignments, and to specify the referents of past terms in modern language unavailable to past scientists, without attributing to them beliefs and abilities that they lacked.

At this point, some clarifications are in order. First of all, Chang (2009a; 2011, p. 419) explicitly takes his motivation from P. W. Bridgman’s operationalism. And as he notes (2009a, §2.1), one frequent criticism of Bridgman’s operationalism is that the operations associated with a term do not exhaust its meaning. Chang’s own response is to agree, and to then show that a focus on operations is nonetheless very useful for understanding various aspects of scientific activity. My response is similar—while operations do not exhaust the meaning of a term, they can often be used to fix its reference.

Secondly, conditions 1 and 2 should be understood somewhat loosely. In condition 1, the set of operations should be understood as capable of changing over time. Moreover, if that set suffices for identifying some entity, it may also be the case that a proper subset of those operations suffices, or even that two distinct proper subsets each suffice on their own. And condition 2 should be understood as not requiring that past and present operations be, strictly speaking, identical to one another. I take it that we can speak of
the sameness of past and present operations, even if, say, today’s laboratory equipment differs from that used in the past, and even if, say, scientists today take more precautions than chemists working in the late eighteenth century.

Thirdly, there’s the question of how to determine which operations past scientists performed. In some cases, reading the texts that they wrote will be sufficient to answer this question. In other cases, we may have to do more. One method that is available to historians and philosophers of science is to attempt to perform such operations ourselves.\footnote{Some examples of historians and philosophers who have relied on this methodology are Jed Buchwald (2008), Hasok Chang (2012, pp. 115–121), and Lawrence Principe (2013).} Success in the laboratory may be good evidence that we’ve successfully determined which operations past scientists performed.

Fourthly, I’ve framed both conditions in terms of whether scientists can use the operations in question. It may be the case that today, scientists no longer perform certain operations because they’ve come up with better ones—ones that are easier to perform, less dangerous, etc. That said, they may still admit that, strictly speaking, performing those past operations would be a way of identifying some entity. This is the sense of ‘can’ that I aim to capture in both conditions.

Fifthly, both conditions make use of the idea that operations can be used to identify an entity $x$. The basic idea here is that scientists often know that some entity $x$ has a number of properties that other entities lack, and they know how to use various operations to determine whether the entity in question has those properties. Hence, the notion of identification to which I’m appealing involves both theoretical and experimental knowledge. To take an example, suppose chemists isolate some gas in the course of attempting to determine the components of some substance. There are subsequent operations that such chemists could perform, e.g., determining whether the gas supports respiration and combustion, measuring its mass, and so on. After performing some such operations, they may be able to conclude that the gas is oxygen, on the grounds that only a sample of oxygen would have those properties. In that case, they’ve identified it. Otherwise, while they may have isolated a sample of oxygen, they haven’t identified it as such.

Sixthly, both conditions make use of present science. The candidates for the referent of a particular theoretical term come from the ontology of current science. And in order to determine which operations are sufficient for identifying some entity, we need to look at current scientific practice, and at
what present-day scientists would say about whether past operations were sufficient for identification. The two conditions are therefore just as fallible regarding the reference of theoretical terms as present science is fallible regarding all that falls within its domain.

Seventhly, and finally, it’s worth commenting about some problem cases. It may be the case that the set of operations associated with some theoretical term from past science contains operations sufficient for identifying some entity \( x \), and operations sufficient for identifying \( y \), where \( x \) and \( y \) are distinct. I take it that in such cases, some account of partial reference is desirable—perhaps one term partially denotes two or more entities, in either McLeish’s (2006) or Hartry Field’s (1973) sense. I take it that another kind of problem case is far more common, namely, the case in which some operations associated with a term from past science can be used to identify some entity \( x \), while others play no role in identifying \( x \). What should we conclude in such cases? My proposal is to privilege the operations that can be used to identify an entity. These operations secure the reference of a theoretical term to that entity. And the operations that turn out not to play a role in identifying the entity represent aspects of scientific activity about which past scientists were mistaken. Since developing a set of operations sufficient for identifying a theoretical entity is no easy attainment, reference will not be too easy. And since reference does not require being right about all of the operations, reference is not too hard.

Now that I’ve clarified my proposal for a non-Whiggish semantic theory, I’ll argue that, when it comes to the example of phlogiston, condition 2 yields the non-Whiggish reference assignment and avoids the Whiggish one. I’ll start with the non-Whiggish reference assignment, according to which ‘phlogiston’ sometimes referred to hydrogen. As I discussed in section 3.1, a number of phlogiston theorists identified phlogiston with inflammable air, which oxygen theorists like Lavoisier identified with his hydrogene. In that case, the operations that such phlogiston theorists associated with ‘phlogiston’ included those operations that they associated with ‘inflammable air,’ and those operations that the oxygen theorists like Lavoisier associated with ‘hydrogene.’ Moreover, these operations survive to the present day—chemists associate them with ‘hydrogen,’ and they can use them to identify hydrogen. Hence, ‘phlogiston’ sometimes referred to hydrogen.

Condition 2 also avoids the Whiggish reference assignment, according to which ‘phlogiston’ sometimes referred to free electrons. As I discussed in section 3.2, the operations that some phlogiston theorists associated with
‘phlogiston’ included those operations by which they identified electricity.\footnote{Perhaps, then, ‘phlogiston’ sometimes partially denoted both electricity and inflammable air. Such a proposal might make sense of Kirwan’s speculation that “phlogiston, in a state perhaps 100 times rarer than inflammable air, … may possibly constitute the electric fluid” (1782, p. 210).} But even if electricity is ultimately made up of electrons, those operations did not amount to operations by which they identified free electrons. The operations that scientists associate with ‘free electrons,’ by means of which they identify free electrons, did not exist until the late nineteenth century at the earliest. Hence, even though we can admit that various phlogistonist explanations of phenomena bear a striking resemblance to the explanations that we currently accept, we cannot conclude that ‘phlogiston’ and ‘free electrons’ co-refer to the same entity.

6 Conclusion

In order to understand past science, we must know, at least to some extent, what past scientists were talking about. In this paper, I’ve proposed one way of determining what they were talking about, which focuses on scientific practices, and in particular, on the operations that past and present scientists perform. I’ve argued that this way of determining what past scientists were talking about is non-Whiggish, in the sense that it doesn’t involve attributing to them any beliefs and abilities that they lacked.

I’m hopeful that the semantic theory I’ve proposed here can be developed and applied in ways that I haven’t discussed in this paper. I take it that it would be desirable to have a theory that says more about those cases in which we conclude that the only non-Whiggish reference assignment is reference failure. If this is all we are told about such cases, I’m not sure that this conduces to our understanding of past science. In this regard, the theory ought to be developed further. Moreover, I’m hopeful that it can be applied to another topic in the philosophy of science that involves the semantics of theoretical terms, namely, the scientific realism debate.

References


Elliott, John (1780). *Philosophical observations on the Senses of vision and hearing; to which are added, a treatise on harmonic sounds, and an essay on combustion and animal heat*. J. Murray, London.


