In the 1950s-60s biochemists searched for a series of high-energy molecules in the cell. Chemists now claim that these molecules do not exist. Yet sixteen distinct claims to have actually isolated or identified them were published. The episode parallels the familiar eighteenth-century case of phlogiston, in illustrating how error is not simply a loss of facts but, instead, must be actively constructed. In addition, the debates surrounding each case demonstrate how revolutionary-scale disagreement is sometimes resolved by differentiating or partitioning empirical domains, rather than by replacement of one theory by another.

The chemical theory is as relevant as phlogiston.

---A biochemist, in the wake of the Ox-Phos Controversy

1. Introduction

On April 20, 1963, noted biochemist Efraim Racker gave a presentation at the annual FEBS meetings. His lab had been chemically dissecting fragments of the mitochondrion, the organelle in the cell most central to energy processing. Racker assessed the status of research in the field with his characteristically dry wit: "Anyone who is not thoroughly confused," he declared, "just does not understand the situation" (Racker and Conover 1963, p. 1088).

Confusion permeated studies especially on one stage of the energy transformations in the mitochondrion: oxidative phosphorylation, or ox-phos. The centerpiece of contemporary theory and practice was a set of unknown high-energy compounds which served as critical intermediates in the transfer of energy. Whoever could successfully isolate and identify these compounds would clearly be in line for a Nobel Prize. But no one had found them. Failures in science may typically recede
gracefully into obscurity, yet another prominent researcher, E. C. Slater, admitted three years later—with no humor intended—that "attempts to isolate compounds with the properties expected of a high-energy intermediate have met with conspicuous non-success" (Slater 1966b, p. 174).

Biochemists failed, textbooks today tell us, because, quite simply, these molecules do not exist. The image of the fictitious molecules, in fact, has become somewhat notorious in the field as epitomizing the failure of the whole research enterprise and its guiding hypothesis. Indeed, one researcher commented disparagingly in the late 1970s that the theory "is as relevant as phlogiston" (Gilbert and Mulkay 1984, p. 32). Here, he alluded to the 18th-century principle of inflammability that was dismissed during the Chemical Revolution and waned in its wake. The biochemist's analogy was particularly apt in two respects, each framing a penetrating puzzle addressed in this paper.²

²Analogies can serve several roles. In this paper, the central analogy between phlogiston in the eighteenth century and the high-energy intermediates of ox-phos in the twentieth century illustrates many of these roles in succession.

First, analogies can be an important historiographic tool for discovery (Darden 1992; Langley et al 1988). But the task can be more sophisticated than simply matching corresponding elements and checking their validity, as Hesse (1966) implies. Boyd (1979), for example, recognizes that an analogy often serves to structure the target domain. That is, not all correspondences in the analogy are obvious; the target domain may be relatively unknown. Sometimes one uses established relationships in the model as a scaffolding or preliminary skeleton—often linguistic—to search for possible relationships in the unfamiliar domain. The analogy is actively constructed.

Second, analogies commonly serve didactically or rhetorically in communicating an unfamiliar domain. In this paper, I use the well known 18th-century case of phlogiston to introduce the less familiar episode about oxidative phosphorylation in the 1950s-60s (§§1,3). The phlogiston analogy allows me, in particular, to highlight two major features of the 20th-century case that have escaped previous analyses of the episode: the transformation of fact to artifact (the construction of error) and the resolution of its deep disagreement.

Sometimes, there is the potential for reverse analogy. That is, where the details of the initial "object" or target of the analogy are better known than the "model," they may serve to understand the original case more fully. The details of the second case may reveal, underscore, or accentuate aspects of the first case. In §10, I show how details of the 20th-century case indeed highlight novelties in the already well studied case of phlogiston. The phlogiston analogy thus functions (at different times and in distinct contexts) in both directions—in each case, developing the analogy further.

Finally, an analogy that is built on or establishes a constellation of similarities can well become a rudimentary, emergent, or nascent general model that embraces the two parallel parts of the analogy. To the degree that similarities exist, one may well postulate or search for an underlying basis for the similarities. At this point, of course, further justification must rest on careful and more robust examination of cases within the appropriate domain. In this paper, I am indeed also concerned about the philosophical resonances between the two cases: how can entities once presumed to exist later be viewed differently, and how is disagreement resolved on such controversial occasions? I would contend—though I certainly do not complete the fuller argument in this more limited paper—that my use of the concept of 'phlogiston' may
First, phlogiston is perhaps the quintessential example of error in science. The biochemist's remark reflected a view that nothing can be worse for a scientist than to believe in a substance that does not exist. Yet some of the world's most outstanding biochemists pursued the ox-phos intermediates vigorously for nearly two decades. Why? Many even published claims to have actually isolated them (see Figure 3 and §6 below). How did researchers justify transforming an accepted claim, once central to research, into "error"? Alternatively, how did fact become artifact? The grand scale of the search for the high-energy intermediates makes it a prime case for understanding the construction of error. Below, I articulate why scientists pursued--and then abandoned--this "20th-century phlogiston" (§§3-6).

The phlogiston analogy also extends to the deep, prolonged debates that surrounded both 18th- and 20th-century cases. As most ox-phos biochemists continued their search, one biochemist introduced an alternative interpretive scheme of energy processing in the cell that did not depend on the elusive intermediates. Disagreement flared (§3; see also Harold 1986; Skulachev 1988). By the time the Ox-Phos Controversy eventually subsided, a revolutionary new consensus of cellular chemistry had emerged (Weber 1991). Given the Kuhnian (1962) style of the reconceptualization, analogous to the more notable Chemical Revolution, the question inevitably arises: how was the disagreement in the ox-phos case ever resolved? Earlier studies of this episode, each adopting a different perspective on science—historical, philosophical, sociological, and native—have not clarified this issue fully. A careful history of the fate of the intermediates, however, helps guide a more complete analysis (§§8-9).

Interpreting how scientists justified their claims of error in this particular episode allows a broader appreciation of the nature of factual claims. Facts fit in a context—not be construed more widely as any entity once presumed to exist, but now relegated to the pile of discarded concepts in science (especially those often labeled as "misleading"). What might we learn by conceiving N-rays, polywater, electric fluid, phrenological analyses, bacterial mesosomes, etc., as "phlogistons"? At this level, the particulars of the original 18th-century case become less significant, serving only as a single exemplar.

Thus, the portrayal of "a 20th-century phlogiston" borrows and extends an analogy in several ways. First, it is a structure, based on a familiar 18th-century case, on which to build a narrative account of the history of the high-energy intermediates of oxidative phosphorylation. Second, it is a historiographic tool of discovery for reconsidering some aspects of the 18th-century case, based on knowledge of the 20th-century case. Finally, it forms a skeletal framework or model for thinking about error philosophically, especially about ontological claims in science.

---

2 An artifact, for experimental scientists, is a result that may at first appear significant, but that ultimately reflects some irrelevant aspect of the experimental procedure, not the phenomenon being investigated. For more on the concept of an 'artifact', see Latour and Woolgar (1979, pp. 60, 174-183), Oxford English Dictionary, definition x), and §7 below.

just of theory or of scientists' social relations alone--but also of other facts. Facts become artifacts, I claim, not by losing their fact-like status, as suggested by Galison (1987), Latour and Woolgar (1979), and others. Rather, "erroneous" facts are actively constructed. They become "new" facts by being repositioned to a different factual background or context.\textsuperscript{4} One may characterize my conclusion using the concept of domain, the bounded constellation of observations or natural phenomena that scientists address as an ensemble in both their theoretical explanations and their experimental manipulations (see §2). A fact becomes an "erroneous" fact when evidence leads us to see it as a fact in another domain (§7).

My discussion about scientists resolving their theoretical disagreement in the ox-phos case also applies more broadly. Even in cases of scientific revolutions, I claim, one conceptual structure does not always wholly replace the other. Instead, one may articulate and redefine the scope of the concept(s), thereby differentiating or partitioning observational or experimental contexts. More precisely, researchers may find evidence for new boundaries to the concepts' respective domains, thereby eliminating contentious overlap (§§9-11).

Finally, in a reflexive mode, I describe how my study adopts a particular methodology and I discuss its implications for science studies (Postscript).

\textbf{2. Domains}

The concept of domain serves as a central vehicle for conveying my claims about constructing error and resolving disagreement—hence, a few preliminary comments (see also Appendix).\textsuperscript{5} Linguistically, we are starved for a term that denotes the

\textsuperscript{4}One may characterize the shift equally well in the sociological language of Latourian actor-actant networks. In Latourian terms, a fact is tied to a certain network of actors, instruments, and phenomena, which are its "allies." When actant-allies "betray" an actor's claims, the network is severed and the claim loses its fact-like status. I claim that an "error" occurs not by severing links, but by building stronger links to other nodes or actants in the network, which are sometimes significantly "remote" from the original links. The network is thereby re-constructed, not merely "de-constructed." Facts are literally re-placed to another network or another place in the same network (see §7).

\textsuperscript{5}The concept of domain has a checkered history. Shapere (1984) introduced the notion of a domain to describe the development of science in separate, relatively independent, self-contained "fields of inquiry" (p. 273), such as electricity, spectroscopy, or rare earth chemistry (p. 279). Elsewhere, he defined a domain as "a body of related information" (p. 276), sometimes as a body of information that posed a problem and "which must, ideally, be accounted for by a theory which resolves that problem" (p. 281). At times he characterized domain as the "organization of subject-matter for study" through personal investigation (1985a, 641). Still elsewhere, a domain is "itself a hypothesis" (1984, p. 281) (all italics mine). While a sympathetic reading might well disentangle and resolve the many apparent discrepancies in meaning, Shapere's notion has generated enough confusion and critical commentary that it is an unstable point of reference. The astute reader will recognize that I borrow elements from Shapere, yet I also depart from his original conception significantly in other respects--
"territory" that a theory is intended to "map." I refer to this specific "territory" as a domain. More formally, a domain is a constellation of phenomena that are closely related--such as might be the focus of an investigation, a series of investigations, or, more broadly, a research lineage. The concept becomes particularly strong when the relationships are causal: in such cases, a domain represents a (postulated or inferred) causal network, relatively isolated (or robustly independent) from other "external" causes. The term domain fits many contexts of reference. The domain of a particular observation would refer to other related observed phenomena (in the same, shared domain); the domain of a concept refers to the observable items we take the concept to map or explain (perhaps causally).

Two aspects of domains are particularly relevant to my analysis. First, the concept of domain highlights the bounded nature of experiments, explanations, concepts, models, or theories. A domain has an explicit or implicit scope, however fuzzy the boundary may be in practice. The concept of domain is both inclusionary and exclusionary: researchers deem some factors relevant, others irrelevant (especially causally). In the ox-phos case, disagreement about domain boundaries was critical.

A second important feature of domains is that they are constructed experimentally and interpreted conceptually. Causal connections, for example, are not self-evident. They must be investigated and often teased apart and demonstrated experimentally. One significant--and often time-consuming--task in science is distinguishing causally relevant from causally irrelevant variables. (What are the necessary and/or sufficient variables for producing a particular phenomenon or effect? In answering this, a researcher articulates a cluster of closely related causes within a larger causal network.) Accordingly, the boundaries or scopes of domains can sometimes become central foci of investigation--especially when scientists disagree.

especially regarding the original impetus for the concept. As detailed in the main text, I focus on the central theme of delineated scope--whether scope of experimentation or scope of explanation. My aim--namely, to clarify scientists' discourse at the level of practice--guides my particular use of the domain concept; one may interpret the concept of domain alternatively, though, from several perspectives on science--experimental, philosophical (explanatory) and social (see Appendix).

One may, of course, conceive causal networks either locally or globally (since no causal network is entirely isolated). Hence, domains may vary in scope, depending on context. Any particular phenomenon, observation, or domain item may thus belong to several nested domains simultaneously (see Shapere 1984, p. 285). The image of nested geographical territories is an appropriate metaphor. Depending on the discussion, a domain may refer to a set of phenomena created in a particular experiment (perhaps with an emphasis on techniques or instruments), to a whole specialty or field of science, or to a set of phenomena of some intermediate scope. In each case, the notion of domain underscores the significance of a boundary separating relevant items from irrelevant ones.

Hence, one may make sense of Shapere's problematic comment that "what counts as a domain is subject to revision or rejection": the items of the domain (the entities or events being investigated) can be regrouped into different domains; the descriptions of those items, and of the domain as a whole,
Domains can thereby, on occasions, be dramatically reconstructed or reinterpreted. Understanding the creative rearrangement of domains was central, I claim, to interpreting the resolution of the Ox-Phos Controversy.

3. The Ox-Phos Controversy

The ideas and methods that fueled the search for the high-energy intermediates of ox-phos emerged seamlessly from earlier biochemical studies of the cell. Ox-phos is part of a much more extensive series of energy-related chemical reaction-pathways in the cell, including the familiar Krebs cycle or citric acid cycle. Together, these reactions channel energy from the chemical bonds in things we eat into a central unit of energy "currency" in the cell: adenosine triphosphate, or ATP. The final set of reactions that yield ATP occur in the inner membrane of the mitochondrion—though the role of the membrane itself was not appreciated when first discovered. During the 1940s and 50s biochemists learned that the energy for ATP is released stepwise in a set of proteins known as the electron transport chain (also the 'respiratory chain' or 'oxidation chain'). (The reader unfamiliar with biochemistry may note that the last step in this chain is precisely where our cells use the oxygen we breathe. As the biochemists noted, these reactions are part of "the secret of life." ) The problem—and it resisted being solved for many years—was: how is energy transferred from the electron transport chain to ATP?

In the early 1950s biochemists viewed energy as transferred in billiard-ball-like fashion from molecule to molecule, like batons in a relay race. The challenge was to trace the path and to isolate and identify the intermediate steps and the enzymes along the way. One could thus reconstitute the system in vitro. In 1953 E. C. Slater made an analogy between ox-phos and another well known reaction and hypothesized an intermediate step or steps. This would involve a yet unidentified molecule or molecules with a high-energy chemical bond (see Figure 1; 'succinyl-CoA', at left, not part of ox-phos proper, served as the explicit analogy; the '~' in 'I~X' denotes the critical high-energy bond; see also Figure 2 for a schematic visualization of the proteins). Slater's 'chemical' hypothesis thus came directly out of and embodied the strengths of the biochemists' tradition. The enzyme-mediated reactions that biochemists knew so well, however, occurred in solution. The ox-phos reactions, by contrast, occurred in the mitochondrial membrane. Ox-phos biochemists thus implicitly extended the interpreted

can be revised; and what it is that is important to investigate about the domain—the problems to be solved regarding the domain—can be reinterpreted. (1985b, p. 8) (see also Kuhn 1970, pp. 24-25, 29; Darden 1991). A closely allied task is interpreting causal categories. That is, scientists reason about how they can generalize or reliably transfer observations about causes from one experimental context to another. The scope of a domain, then, also establishes indirectly how a researcher may expect to generalize his or her conclusions—another potential point of disagreement. While not speaking explicitly in terms of domains, Knorr-Cetina (1981), Latour (1987), and Myers (1990), among others, have all emphasized both the interpretive and potentially contentious aspect of generalization. The role of ascertaining domains thus has far-reaching consequences for how scientists reason and argue.
domain of their enzymatic concepts by addressing slightly unfamiliar cases. But they did so confidently, based on their previous successes. For the next two decades, the search for Slater’s high-energy intermediates would be the central task of the field (Slater 1966c, p. 382; Racker 1968, p. 32; Skulachev 1988, p. 399).

One may contrast the chemical view with the one we now use. Our current perspective is wholly reconstructed, as profoundly as Lavoisier’s was from the notion of phlogiston--but perhaps even more subtly and elegantly. We now see an intermediate energy state, but not one of high-energy bonds. Rather, the electron transport chain creates a pH difference and an electrochemical gradient across the mitochondrial membrane--a form of energy more analogous to a reservoir or battery. Peter Mitchell, who eventually received the 1978 Nobel Prize for his work, first presented this view in 1961 (Mitchell 1961). Of course, his view, too, emerged seamlessly from his own background. Mitchell addressed problems of "active transport": how proteins use ATP to move molecules across membranes, particularly where they might concentrate inside a cell. Mitchell also noticed how the reaction could be reversed: a cell might use a concentration gradient to energize ATP. Mitchell stressed that chemical reactions could be linked to osmotic movement--and he dubbed his notion the 'chemiosmotic' hypothesis. Mitchell's scheme was able to account for a number of observations that remained somewhat puzzling in the chemical views of ox-phos, such as the need for closed membranes.

From the current vantage point of membrane potentials, the hypothesized high-energy chemical intermediates do not--and never did--exist. They are "a 20th-century phlogiston." Like the earlier debate over phlogiston, this episode, too, generated deep-rooted tensions about claims of existence. In ox-phos, over nearly two decades, as one researcher described it, several models, "each in several versions, were vigorously promoted and roundly condemned. . . . This was a time of strife, dominated by controversy over the essential nature of energy coupling whose flavor was at times almost Byzantine" (Harold 1986). Yet another text writer remarked on the "contentious, often rancorous discussion" (McGilvery and Goldstein 1979, p. 390). Slater himself noted regretfully that the 1960s was "not one of the happiest periods in the history of mitochondrial research" (1981, p. 29).

The ultimate adoption of Mitchell’s concepts was revolutionary, both conceptually and experimentally. First, the chemiosmotic mechanism relies critically on a membrane. Without the structural integrity of a barrier separating inside and outside compartments, the concentration gradient will "leak" and the energy dissipate. Further, the electron transport chain must be embedded in the membrane in a specific orientation, so that it generates a gradient in one direction. The enzyme that forms ATP, likewise, must be oriented in the opposite direction. Direction--inside and outside, orientation in space--matters chemically. To someone accustomed to picturing chemical reactions in terms of something like the bouncing ping pong ball model of gases, the spatial aspect of these processes as proposed by Mitchell were strange indeed. The chemiosmotic formulation implied experimentally, for instance, that one could not--as biochemists like Racker and Slater generally did--simply tear apart the mitochondrion, isolate its essential components, throw them back together again in a
test tube and expect them to work. The attention to direction and ion gradients—and thus to membranes and the movement of reactants—required a whole new approach to biological chemistry, one that is still being developed. Mitchell’s conception of ox-phos required a shift from scalar chemistry (based on magnitude alone) to vectorial chemistry (based on magnitude and direction together). Abandoning the search for the high-energy intermediates, and following instead a new experimental gestalt, marks in part the revolution that has established bioenergetics as a new, fairly distinct field of inquiry (for further analysis of how Mitchell’s ideas differed from his predecessors, see Weber 1991). The sense of revolution merely underscores the question: why was there a rear guard at all? How could biochemists have sustained belief in the high-energy intermediates?

4. Triangulating and Intervening

First, biochemists in the 1960s were able to reason empirically about the presence of the intermediates even though none of them had yet been identified or isolated. Sometimes they could even specify properties of the intermediates. The biochemists’ conclusions, though largely circumstantial, were triangulated from numerous experiments and even proved effective in predicting and guiding laboratory interventions. Laboratory practice supported their beliefs.

For example, Slater’s original hypothesis had been inspired in part by the action of a set of chemicals that could ‘uncouple’ oxidation and phosphorylation. When these chemicals were present in the cell, the energy of the electron transport chain would be used, but no ATP produced. Indeed, the ATP reaction would itself reverse. This implied an intermediate step where the uncoupling chemicals would function. Uncoupling did not indicate what the intermediate step was: but from known points of reference, simply that one must exist.

Another line of evidence was based on tracking radioactive atoms from molecule to molecule through reactions. These experiments exploit the fact that in chemical reactions, atoms or chemical groups are transferred from one molecule to another. This exchange can be traced by labeling a compound with a specific radioactive atom and then identifying the molecule where the radioactivity later appears. These ‘exchange reactions’ essentially allow biochemists to discover unknown reactants at earlier or later steps than ones that are already known. One can chemically “dissect” a series of reactions initially known only by its endpoints.

One important exchange reaction was discovered between ATP and ADP. ATP is formed by combining ADP, adenosine diphosphate, with a phosphate. The ATP-ADP exchange reaction indicated that in the last step, ATP was produced from a separate ADP molecule (Lehniger et al 1958; Wadkins and Lehninger 1958). Since the ADP alone was not energized, the energy for the ATP bond must have come from another source: namely, along with the phosphate in a ‘high-energy phosphate bond’. This was plausible, as such high-energy phosphate bonds had been identified at several other steps in the cell’s energy reactions. Thus there was not merely some circumstantial hint of a high-energy intermediate. There was evidence that it contained, more specifically, phosphate.
Another, even more dramatic form of evidence involved actually predicting the behavior of the high-energy intermediates and controlling them in the laboratory. According to Hacking (1983), Galison (1987) and others, these "interventions" into nature are the substantive form of proof in science, what makes things "real." If the intermediates were real, for example, one might be able to use them to fuel other energy-requiring reactions in the cell. Slater had suggested as much, though informally, in his original paper. This application of the intermediates' energy was, in fact, done.

The strategy was simple: block energy flow along certain known pathways and essentially force it to flow through what could only be the intermediates. The challenge, of course, was to determine how this could be done experimentally. Lars Ernster devised a method in 1963 (Ernster 1963). He diverted energy from the last section of the electron transport chain to the first part (refer to Figure 1). He first inhibited energy flow in the middle section of the electron transport chain (upstream of cytochrome b). Then he added a source of energy "downstream." With another inhibitor (oligomycin) as a further "roadblock" (between I–X and ATP), he prevented the production of ATP or the use of ATP as a source of energy itself. With the assumption that the three intermediates (I–X in Figure 1) were the same or could exchange energy, the energy had only one place to go. The high-energy intermediates redirected energy to the first part of the chain. There, the energy induced electrons to flow in reverse (from f0 to DPN). That is, rather than move down energy levels, as they normally did, they went up energy levels, creating a measured accumulation of DPN. The reversed electron flow was striking--and like nothing that occurred naturally. Ernster had used the conceptual model to predict a complex effect not found in nature and to design the experimental conditions to produce it. Ox-phos biochemists had found a way to manipulate their proposed energy state, to intervene in nature.8

Together, the interpretations of uncouplers, the ATP-ADP exchange reaction and the reversal of electron transport show why biochemists were confident experimentally that the high-energy intermediates existed. One should note, in addition, that these three examples come from a larger set of arguments that were well-rehearsed in texts and occasionally in review articles—even as late as four years after the Nobel Prize had been awarded to Mitchell (see Tzagoloff 1982). The experimental foundation for the chemical hypothesis was quite rich. Thus there was no irony intended when one researcher boldly claimed in 1965 that study in ox-phos had advanced beyond the phase in which the intermediates were merely hypothetical (Griffiths 1965, p. 116).

Through the privilege of retrospect, we may recognize that all the evidence amassed for a high-energy intermediate molecule was only sufficient for concluding that there was an intermediate energy step of some kind. But while biochemists might have been able to differentiate these claims conceptually at this time, in practice they did not. Indeed, interpreting these two claims as essentially the same had led to their

---

8In much the same way, of course, 18th-century chemists had been able to juggle phlogiston from one source to another (see below, §10, on "Reassessing Phlogiston").
great success in tracing related biochemical pathways and ATP synthesis elsewhere in the cell. Coming to accept the unsuspected differentiation between energy step and energy molecule was much of what this episode was about (see §9).

5. Interpreting "Falsifying" Anomalies

The reasoning discussed above was all developed from principles and methods available within the chemical paradigm or research lineage. There was disagreement, to be sure, but none that challenged the basic assumption that intermediate molecules existed. What is equally if not more important, though, is the reasoning that was advanced in the face of criticism from the framework or paradigm we now accept. Under many schemes of conceptual change in science, important judgments are only made in the contexts of alternative explanations (e.g., Lakatos 1970; Bayesian models; Latour 1987). In this case, biochemists continued to pursue the intermediates even after the current theory was introduced. Their response to criticism reveals even more deeply how they reasoned about our 20th-century phlogiston.

Perhaps the most penetrating criticism against the chemical hypothesis was that despite years of intensive effort, the intermediates had never been found. As phrased by Mitchell, they were "elusive to identification"--an understatement that merely underscored the experimental thorn (1961, p. 145). This failure could easily be accounted for, Mitchell gently urged, if one admitted that the ox-phos intermediates simply did not exist. Further, the "conspicuous non-success"--to borrow Slater's phrase--was paralleled with the complementary failure to reproduce these reactions in vitro independently of an intact membrane. The reason, according to the chemiosmotic view, was the absolute requirement of a topologically closed membrane in maintaining a gradient. The observational correlation was taken to mean that membranes (and hence membrane gradients, rather than chemical intermediates) were causally linked to the domain of ox-phos. From this perspective, the paired experimental deficits were telling. They were anomalies that threatened the very fabric of the chemical approach--falsifying its basic premises.

For the biochemists, however, Mitchell's criticisms were hardly more than familiar observations, and certainly not an epistemic challenge. Biochemists, like Slater, were well aware of the problem and its magnitude. One textbook pined that "even if we do not have most of the information necessary for further reasoning, it is unsatisfactory to have blind faith in a mystical high-energy state. Biochemistry is supposed to be on a sounder basis than parapsychology" (McGilvery and Goldstein 1979, p. 391). Yet from the chemical view, the experimental problems about isolation and the membrane did not cast doubt on the well-established empirical knowledge about the intermediates. Rather, they revealed important and unanticipated facts about the structure of proteins in membranes. This, in turn, demonstrated the need for new experimental techniques.

Biochemists were accustomed to studying enzymes that functioned in an aqueous or water environment. As noted above, however, the electron transport chain is located in the mitochondrial membrane--a medium of lipids or oil-like molecules. As the saying goes, oil and water do not mix: their chemistries are quite different. Thus,
biochemists found, much to their dismay (and later frustration), that one could not easily extract the proteins in solution. Biochemists had to cope with the proteins of the respiratory chain still buried in miniature vesicles or 'submitochondrial particles'. The recalcitrance of the system to simple decomposition was notoriously bothersome. Albert Lehninger (now renowned for his classic biochemistry text) noted, for instance, that "it was part of the biochemical Zeitgeist that particles were a nuisance and stood in the way of purification of the respiratory enzymes" (Lehninger 1964, p. 6). This applied as well to the effort to isolate the high-energy intermediates. While Mitchell was trying to focus on the membrane as functionally integral; other biochemists were trying to get rid of is as superfluous and interfering. Ironically, the inadequacy of the experimental techniques more than adequately accounted for the biochemists' persistent troubles.

Elsewhere, however, Lehninger recognized how the inability to successfully isolate functional proteins from the membrane might be viewed more positively. "There may be," he suggested, "a biological necessity for structural organization of these catalysts in a moderately rigid, geometrically organized constellation in the membrane." This would, for example, "minimize the path distance between slowly diffusing large molecules and . . . maximize probability of interaction" (1960, p. 952; see also Chance's image, Figure 2). The membrane may have had a skeletal role that contributed to the enzymes' interaction—an unprecedented but hardly unlikely circumstance (see also Grabe 1958). The frustration in dealing with the experimental system, rather than showing how the assumptions of the 'chemical' hypothesis were wrong, apparently revealed an unanticipated dimension of biological organization. The "moral" for biochemists was that they needed to search more creatively for ways to isolate or prepare such complex—and apparently quite fragile—membrane-bound systems.

Like Lehninger, Efraim Racker was impressed by the highly organized assembly of ox-phos components, many of which his lab had identified. At the same meetings in 1963 where he had depicted the confusion in the field, Racker, too, considered how the elaborate organization might be linked to a chemical mechanism. He refrained from saying more about the structure, however, because it was too complex. He apologized: "I shall not show you a scheme of the topography of the various factors in mitochondria because I promised to keep this presentation simple. But," he offered in consolation, "I carry a picture of it in my wallet, together with photos of my wife and daughter and I'll be glad to show all three of them to anyone who cares to see them" (Racker and Conover 1963, p. 1091). The humor in Racker's comments was not gratuitous: it underscored to the "family" of ox-phos biochemists the familiar complexity that justified both why the intermediates had not been isolated and why the whole system had not been reconstituted in vitro.

From Mitchell's perspective using the chemiosmotic hypothesis, the laboratory experience of needing a membrane was falsifying evidence against the chemical hypothesis. Few biochemists denied that Mitchell offered a plausible explanation. But while acknowledging the very same fact, they regarded it as evidence for organizational complexity. Biochemists could justify their 20th-century phlogiston, even with an alternative explanation available.

If Mitchell had identified an anomaly in his 1961 critique, it was certainly an ambiguous one. But one should note more fundamentally how the interpretations could
differ. The contrary assessments were grounded, not so much in different theoretical commitments, as in separate observational or experimental contexts. The same fact connected to different primary points of reference (crudely, membranes on the one hand, and proteins on the other). Using those different points of reference, one tended to draw the boundaries of the causally relevant factors in the ox-phos domain differently (see Figure 5a). The differently interpreted domain, marking divergent orientations, offers a central clue to understanding the history of the high-energy intermediates of ox-phos. But the story is still incomplete without considering a third set of claims supporting the intermediates.

6. Eureka?

The case of ox-phos would not be nearly so interesting if no one had ever claimed to have actually isolated or identified the high-energy intermediates. But somebody did. Nor would it be so interesting if such a claim turned out to be fraudulent--which also happened (twice). What makes the ox-phos episode so fascinating is the fact that not just one, or two, but at least sixteen different claims were published, each offering evidence for a separate intermediate molecule or isolate (Figure 3). Further, most such claims were presented not just once, but in a series of successive papers or research reports. In its multiplicity of published existence claims, the ox-phos case is even more philosophically engaging than the history of phlogiston.

None of the intermediate claims, however, ever survived in the long run. Indeed, it was difficult not to notice the regularity with which claims were first proposed and subsequently abandoned. One text writer noted the confusing lesson for his readers: "no worse fate could befall anyone working on oxidative phosphorylation than to solve it" (Tzagoloff 1982, p. 131).

One may well imagine that a creative scientist could finesse the general concept of an intermediate. But if no intermediate existed, how could one support specific claims empirically? Here, the claims had shifted from hypothetical intermediates to real intermediates, from high-energy intermediate steps to high-energy intermediate molecules. Also--and equally important to understanding the historical errors--why did no claim endure?

A number of claims were suggestive, at best. For example, Jack Purvis, a post-doc in Slater's lab, isolated a form of a familiar molecule that was not involved in electron transport and that accumulated under conditions one would expect for a high-energy intermediate (Purvis 1958; 1961; Purvis and Lowenstein 1961). It produced ATP when ADP and phosphate was added. Other biochemists, though, did not trust Purvis' measuring methods and could not repeat his results (e.g., Klingenberg and Bucher 1960). Investigations were later dropped when a graduate student in the same lab, Karel van Dam, found that the amount of intermediate energy suggested by Purvis's results exceeded what the cell could produce. Purvis' intermediate began as a promising indication, but did not exhibit all the properties expected of its role.

Another suggestive claim was introduced by Ronald Butow and Efraim Racker (1965a; 1965b; Butow 1965). They were looking at factors that controlled the flow of electrons through the electron transport chain. One agent, \(\alpha\)-phenanthroline, inhibited
electron flow. It could also bind to metals (as a 'chelator'). Iron was part of a protein complex in the mitochondrial membrane, but it was not part of the electron transport chain proper. Butow and Racker assumed that their agent would bind to the iron, and reasoned that it could affect electron flow if the iron was itself bound to an electron carrier. Non-heme iron, they reported, should possibly be seen "as a component of the non-phosphorylated high-energy intermediate of oxidative phosphorylation" (1965a, p. 160). Racker's lab later found this proposal consonant with findings about a system similar to ox-phos in yeast (Schatz and Racker 1966). Still later, Racker, working with another biochemist, successfully reconstituted a fully functional electron transport chain without iron, confirming that iron was not part of the chain (Yamashita and Racker 1968; 1969). In referring to the earlier work, however, their claims were less definitive: iron was either part of the phosphorylation mechanism or--reflecting Lehninger's earlier comments--had a structural role in the membrane (1969, p. 1226). The role of iron was suggestive because all the evidence was indirect: based more on the absence of iron than on any particular effect characterized in its presence. For example, no iron-containing compound was isolated that could produce ATP.

Paul Boyer expressed his doubts, suggesting that a reasonable alternative was that the agent, o-phenanthroline, might affect the process in a way other than by binding to metals such as iron (Boyer et al. 1966). A group in Japan later confirmed this suspicion. They showed that the agent could inhibit electron flow even in iron-deficient fragments. Further, they showed that m-phenanthroline--a close chemical variant that had no metal-binding ability--could also inhibit electron flow (Imai et al. 1968). While biochemists continued to see iron as part of the ox-phos system more generally (e.g., Baltschefsky and Baltschefsky 1974), they could reject its possible role as an intermediate. The suggestive effects that Butow and Racker had first observed were indeed "effects," but they were eventually attributed to causes other than the presence of non-heme iron.

For a number of proposed intermediates, ox-phos biochemists drew on theoretical models. Chemists often searched the range of possible reaction mechanisms for potential 'model reactions' (see, e.g., Bechtel 1988). The vicarious search offered likely candidates and thus more concrete clues for the experimentalist. At one point, Lardy and Ferguson (1969) could list fifteen such mechanisms under consideration. Often, these "claims" became stronger in the eyes of others than for their original proponents (see Figure 4).

One might easily dismiss the theoretical proposals as speculative only. Many claims, however, were supplemented with more concrete demonstrations. Thus, Youssef Hatefi (1963) revived his earlier proposal that "coenzyme Q may occur in a high-energy state (QH$_2$-I)" when others showed that quinol phosphate could phosphorylate adenosine monophosphate (the precursor to ADP; see Hatefi and Quiros-Perez 1959). His conceptual picture, once floating exclusively in the realm of indirect reasoning, had been given an experimental anchor. Hatefi coupled the theoretical map to an observational benchmark.

Theory and experiment were combined a similar way for several proposed intermediates, some related to molecules already known in the ox-phos reactions. Jui Wang implicated a ferrohemochrome imidazole; Arnold Brodie, a naphthoquinol
phosphate; and Audrey Painter and Edmund Hunter, a glutathione (Brinigar, Knaff and Wang 1967; Cross, Cross and Wang 1970; Wang 1970; 1973; Brodie and Russell 1961; Russell and Brodie 1961; Brodie and Watanabe 1966; Watanabe and Brodie 1966; Painter and Hunter 1970a; 1970b). First, they each demonstrated that the molecules could phosphorylate ATP \textit{in vitro}. Reactions involving these molecules could occur in test tubes, therefore, not merely on paper. Brodie's reaction was also destroyed by one of the uncouplers. Further, a natural enzyme (found in bacteria) could form the compound. Finally, Painter and Hunter's results gained considerable profile because the reactions seemed to occur in a membrane-free extract. Though all these claims were originally based on model reactions, they were less circumstantial due to experimental work.

But in all three cases, no one could confirm the specific presence of the appropriate compounds or reactions in the cells themselves. Racker, who like many others found Painter and Hunter's findings promising, later admitted that they were simply "not reproducible." More telling, they were not reproducible "in several laboratories" (Racker and Horstman 1972). Brodie's claims slipped into obscurity as no further evidence of a natural quinol phosphate materialized. Wang presented some of his growing set of data at a conference in 1971, but by that time biochemists in ox-phos were accustomed to the flaws of numerous intermediate claims and were highly critical (Wang 1973, pp. 766-68). In the discussion that followed, Racker and Hager each suggested that Wang's "intermediate," detected only through radioactive phosphate, had received its phosphate via already produced ATP. Richard Cross, who had done Wang's experimental work, isolated the source later, after moving to another lab. The radioactive phosphate was not attached to a protein, as both he and Wang had earlier supposed, but rather to ATP (Cross and Boyer 1973). They had not isolated an intermediate but--somewhat anti-climactically--ATP itself. The several claims based on model reactions were experimentally leading. But as partially substantiated or "fuzzy" possibilities, they never matured as researchers anticipated. The proposed compounds performed certain functions, but not in the appropriate causal context or domain that would locate them in the cell's ox-phos reactions.

For other claims, however, the evidence was more direct still. In 1960, for example, Gifford Pinchot extracted a component, NAD\textsubscript{~E}, from a molecular complex in the bacterium \textit{Alcaligenes faecalis}. He showed its ability to produce ATP in two ways (1960, pp. 929, 937). Three alternative explanations had been "examined and rejected." "It is therefore proposed," Pinchot concluded confidently, "that this compound is a high energy intermediate of oxidative phosphorylation" (p. 929). In periodic articles over the next eight years, Pinchot and his colleagues addressed further aspects of these reactions: the roles of magnesium ions and pH (Scocca and Pinchot 1963; 1968), the presence of both an intermediate step (Pinchot and Hormanski 1962) and of a second transfer enzyme (Pinchot, Hormanski and Scocca 1964), the function of the intermediate as an ETC inhibitor (Pinchot 1965), and a more general elucidation of the steps (Pinchot 1963). Pinchot had developed a repeatable and at least moderately robust phenomenon.

Pinchot's work, however, was ultimately not regarded as evidence for an intermediate--and it is pivotal to understand why. Pandit-Hovenkamp (1965) criticized
Pinchot's findings on the basis of pH and on how ATP was measured. She argued that, like the model reactions, Pinchot's extract could not fill all the causal roles expected of it. Pinchot himself commented at one point that his original announcement, though "compatible with the hypothesis that an energy-rich soluble intermediate had been isolated," did not exclude other "less exciting explanations" (1963, p. 1077). Griffiths (1965) explained though--and this is critical--that Pinchot had indeed isolated something. But it was an NAD-coupling factor, an element necessary for the ox-phos reactions, not strictly a high-energy intermediate.

Pinchot's experimental results were thus not wholly invalidated. He had produced ATP (repeatably) given the procedures he outlined. This was not error. But the procedures did not fit in the causal network (experimentally, conceptually) where the intermediate should be. The error was in where he situated his conclusions or generalizations. Coupling factors were important to ox-phos, but not as centrally as the high-energy intermediates. The claims were legitimate, but in a different--and far less interesting--domain.

This pattern of "repositioning" experimental outcomes in a different, less central causal context was echoed in several other cases. Rao Sanadi's "factor B," proposed over a five-year time span (Sanadi et al. 1963; Lam, Warshaw and Sanadi 1967; Sanadi, Lam and Kurup 1968), was, like Pinchot's, later dismissed as an intermediate. Its behavior closely matched a coupling factor, $F_3$, already isolated by Racker (Lardy and Ferguson 1969, pp. 999-1003; Sanadi and Joshi 1979). Again, the proposed intermediate did contribute to producing ATP, but not via the causal path as first interpreted. David Griffiths' data on NADH~P (1963; Griffiths and Chaplain 1962a; 1962b) was likewise reassigned to another peripheral enzyme. This, too, was a part of ox-phos, and only an "error" when viewed as a high-energy intermediate.

Perhaps the most promising and well received claim was Paul Boyer's proposal for phosphohistidine (1963; Boyer et al. 1963; Boyer et al. 1964; Lindberg, Bieber, et al. 1964). His lab was well respected and the original proposal, appearing in Science, was well profiled. Phosphohistidine was present in a protein that produced ATP and, best of all, it was a soluble system. The "intermediate" was also one of the most short-lived. Within a few months, phosphohistidine had been traced in *E. coli* as well. But there it was part of the succinyl thiokinase reaction (Kriel and Boyer 1964). That seemed an exciting clue about the possible ubiquity of phosphohistidine in phosphorylation reactions. Unfortunately, it also signaled a need to double check the results in the mitochondrion. The results that had offered so much hope were soon isolated to the succinyl thiokinase reaction, also in the mitochondrion--still an important energy reaction (related to the Krebs cycle), but not part of oxidative phosphorylation (Mitchell, Butler and Boyer 1964; Pressman 1964; Slater and Kemp 1964; Slater, Kemp and Tager 1964; Bieber and Boyer 1966; Ernster and Schatz 1981). Not long after, phosphohistidine was also found as a component in a second enzyme--nucleoside diphosphokinase--supporting its role in phosphorylating reactions but, again, not those of ox-phos (Norman, Wedding and Black 1965). Boyer and his colleagues had not erred in proposing a role for phosphohistidine (see Lindberg, Duffy, et al. 1964). They had erred in where they had placed it on the causal map of energy reactions in the cell.

In the wake of the enthusiasm and then disappointment regarding
phosphohistidine, a proposal for phosphoiodohistidine was not likely to engender immediate support. But, again, the shifting horizon of evidence is telling. When Perlgut and Wainio (1966a) suggested phosphoiodohistidine as an intermediate, they had verified its presence in the cell and shown that it fit many of the criteria for an intermediate discussed in the literature (e.g., Griffiths 1963). More importantly, perhaps, they were well aware of the fate of Boyer's claims and explicitly addressed problems which had been raised earlier—namely, the rates at which the intermediate versus ATP were phosphorylated (the same problem that had plagued Pinchot and would later plague Wang). Answering further criticism, Perlgut and Wainio (1966b) used chromatography to give a "more positive identity" to their unknown compound, by comparing it with synthetic phosphoiodohistidine and distinguishing it from other iodine-related compounds previously identified in the mitochondria.

Holloway et al. (1967a) announced that using the reported procedures, they were able to successfully reproduce Perlgut and Wainio's results. But—lest one construe mere reproducibility as a form of support—they did not reach the same conclusions. They noted that the method used to identify the iodohistidine was based on a color reaction that was non-specific. They identified seven possible compounds in the replicated extract, including four known phospholipids, that could produce the same color reaction. All, however, they showed were unstable when treated with alkali—unlike iodohistidine, if it had been present. They further noted that Perlgut and Wainio themselves admitted that they had failed to isolate the iodohistidine by secondary means. Iodohistidine, they concluded, contrary to the original claims, was not present in mitochondria. What had Perlgut and Wainio observed, then? In a further study, the same team of critics (1967b) showed how iodide itself, which had been used in the incubation procedure, if not sufficiently washed out, could account for all the observations. Perlgut and Wainio's results were quite factual. They were simply different facts than the ones they first reported. Of course, the ultimate conclusions here were downright trivial. Due to conservation of matter, the iodide that was added (and not removed in subsequent procedures) remained as iodide. That was hardly novel, though it was certainly supported by Perlgut and Wainio's results. The only "error" had been in attributing the original observations to the domain of ox-phos.

In these five cases—NAD~E, Factor B, NADH~P, phosphohistidine and phosphoiodohistidine—the results lost significance in the expected domain or causal network by gaining significance in another domain. Local causal connections remained the same but they became connected to the larger network along different lines. The "facts," as originally perceived, were not false. But they did lose their intended status as they moved to another observational context. Evidence for these intermediates became irrelevant to ox-phos by becoming relevant to another causal process.

One may summarize the fate of the high-energy intermediates of ox-phos, then, in two general ways: (a) researchers could not generalize their findings from one set of experiments to a wider domain or scope of phenomena, or (b) results interpreted as fitting within one domain or causal network were later found to fit in a different domain. In the former, prospective facts were discarded; in the later, the facts were displaced.

7. Fact, Artifact, and Other Fact
The claims for our 20th-century phlogiston add an important and underappreciated dimension to the meaning of 'fact', 'error', and 'artifact'. First, the ability (or inability) to produce a stable and replicable 'effect' has been accorded varying significance by different interpreters (Latour and Woolgar 1979; Hacking 1983; Galison 1987; versus Collins 1985). Reproducibility was clearly central for many cases in ox-phos. Painter and Hunter's membrane-free reactions and Purvis's 'extra-NAD' were both initially dismissed because no one could replicate the results. Even a fraudulent claim--Webster's RCCF (see Figure 4)--was rejected, well before its fraudulent nature was revealed, by those who could not confirm the provocative claims in their own labs (Sanadi 1965; Schatz and Racker 1966; Griffiths 1965). Findings that could not become a resource for further study had no standing (Knorr-Cetina 1984; Latour 1987, pp. 119-21; Hull 1988, esp. Chaps. 8, 10).

But the critique of phosphoiodohistidine indicates that replication is not everything. One can repeat results and still challenge their relevance. As molecular biologist Walter Gilbert once cautioned, sometimes "you can reproduce artifacts very, very well" (Judson 1981, p. 170). The error in phosphoiodohistidine--at least as an intermediate--was that one could not connect the experimental results to the appropriate causal context or domain. In the same way, the suggestive findings regarding the model reactions, the data on phosphohistidine or non-heme iron--though one could reproduce them all--never settled into their intended causal roles in ox-phos. Replicability (in some practical form) seems essential for incorporating findings into an expanding experimental network, but it does not enable a researcher to distinguish meaningful from useless claims.

The dichotomy between true and false, right and wrong, reliable claim and error, has been recast by some as the distinction between fact and artifact (e.g., Latour and Woolgar 1979; Latour 1987; Hacking 1983; Galison 1987; Franklin 1990). Scientists strive to "carve away the background" of "artificial" experimental effects to see nature. Interpreted in this way, every proposed intermediate of ox-phos was an artifact: we were not seeing nature, but an incidental and meaningless residue of the experimental design.

But in the case of our 20th-century phlogiston, Boyer, Pinchot, Sanadi, Griffiths, Butow, etc., did view nature. The presumed intermediates did not become artifacts. Rather, they became other facts. Boyer's phosphohistidine was a fact: it was part of the succinyl thiokinase reaction. Pinchot and Sanadi had each isolated something essential: a coupling factor. Griffiths had revealed the NAD-NADP transhydrogenase reaction. Butow and Racker had indeed found a way to inhibit electron flow, simply not via iron. All expressed nature. We simply had to determine precisely which corner of nature we were viewing: how was each phenomenon that resembled an intermediate related to the larger causal network? Where were these "facts"? The role of further experiment, then, was to find how the processes surrounding each proposed intermediate were situated in a broader domain.

In each case, error was itself a fact. Biochemists knew that there was an error and knew exactly what the error was. They rarely concluded simply that the proposed compound did not exhibit the properties expected of a high-energy intermediate. Rather, each erroneous intermediate had a specific alternative meaning. It gained its
other meaning by being tied to another domain. Indeed, identifying the other domain—whether it was the succinyl thiokinase reaction (Boyer) or NAD-coupling (Pinchot)—was integral to the biochemists’ reasoning. They concluded that each proposed reaction or compound was not a high-energy intermediate of ox-phos by identifying its role elsewhere. Reliable facts are robust (Wimsatt 1981). But so, too, are reliable errors (Culp 1994). Wang’s "discovery" of ATP and Perl gut and Wainio’s isolation of iodide(!) certainly epitomized unremarkable facts. They were, nonetheless, facts.\(^9\)

Researchers in ox-phos did, of course, make significant distinctions between claims. Some claims were causally relevant to ox-phos; others were not. A claim about an "erroneous" high-energy intermediate was thus a claim about its (ir)relevance to the particular domain of energy transfer between the electron transport chain and ATP. 'Artifacts' were those that could seem causally relevant given some information, but lost this status as causal relationships became better known. Relevance—and any claim about fact or error—was thus domain-dependent. Viewed from the domain of ox-phos, for example, the effects attributed to phosphohistidine were an artifact. Viewed from the domain of the succinyl thiokinase reaction, however, those same effects reflected a (true) fact. Conclusions followed from adopting a certain perspective, though not in the sense of a theoretical commitment or paradigm-induced perception. Perspectives were positional, based on adopting an experimental or observational standpoint or domain as primary.

In brief, all the documented claims of high-energy intermediates were facts. Viewing them specifically as artifact or error involved assuming that the causal domain of ox-phos was primary. Without positioning one’s perspective within a certain domain, there are no artifacts—only facts and other facts. Articulating the broader domain of each proposed intermediate experimentally was critical. In this way, ox-phos biochemists sorted relevant from irrelevant—or acceptable fact from "error." Eventually, they found each intermediate to be an irrelevant fact—one that could also masquerade in certain experimental contexts as causally relevant.

8. Shifting Research Horizons

In the late 1960s, ox-phos researchers were virtually as confused as they were in 1963. Racker’s comments again expressed the status of the field. "Nature may be difficult, but she is never malicious," he quoted Einstein as saying. He then added that "Einstein had obviously never worked on oxidative phosphorylation" (see Rowen 1986; Lardy and Ferguson 1969, p. 991).

Despite the inability to isolate a high-energy intermediate, biochemists always seemed able to account for their failures. The intermediates could be present in small quantities; one could well expect them to be unstable or short-lived; etc. (e.g., Griffiths 1963; 1965; Chance et al 1967; Greville 1969; Racker 1970). The failure to solve the isolation puzzle reflected more on the ingenuity of the experimenters or the limitations

\(^9\)Framing relevant fact and error symmetrically in this way has important implications, which I discuss further in the Postscript.
of available technology than on the concept itself (see Kuhn 1970, pp. 35, 37, 80; Donovan, Laudan and Laudan 1988, pp. 21-26, 377). One must not miss, therefore, the implicit meaning in Bieber and Boyer's apparently casual comment that "proof that something does not exist is obviously difficult to attain" (1966, p. 5382).

The search for the intermediates did wane, though. This was largely because no one could perpetuate a claim in the relevant experimental network. Support, in the form of continued use or pursuit, relied on productivity and actual performance (see also Robinson 1984). Biochemists did not respond so much to criticism about the intermediates as to their own frustration and disillusionment. They gradually exhausted the technical possibilities they could imagine and--without acknowledging their action explicitly--abandoned the search for the high-energy intermediates in lieu of other more tractable projects. They generally did so, however, without also formally rejecting the promise of finding an intermediate by some yet unknown method. Biochemists were thus eager to hear reports of a possible intermediate as late as 1975, though the claims turned out to be fraudulent (see oleoyl phosphate in Figure 4). "Everything remains possible in oxidative phosphorylation--" Slater once remarked, "except the easy solution" (1966a, p. 542).

9. Differentiating Domains

The history of the high-energy intermediates is more than a case study in error. It offers clues for interpreting the broader controversy, in which the search for the "20th-century phlogiston" played a major role. Pinchot's, Sanadi's, Boyer's and others' claims each became artifacts by being transformed into other facts. That is, they were relevant and "correct" in a substantially different causal context or domain. The resolution of debate in ox-phos likewise becomes clear when one views it in terms of the changes in domains or the scope of each hypothesis.

Originally, the chemical hypothesis was the prevailing "map" of the ox-phos domain. When Mitchell introduced the chemiosmotic hypothesis, he offered a substantially different map (5a). The maps differed in what they portrayed as the relevant causal factors--that is, in where they located the boundaries of the ox-phos domain. These two versions of the domain overlapped significantly, however, thereby generating substantial debate--ostensibly about which hypothesis was correct. But how did the controversy itself ultimately end, and what was the resulting status of the domain of ox-phos?
First, the sense in which a high-energy bond or chemical intermediate is a central causal link between the electron transport chain and ATP is no longer supported. A number of dramatic demonstrations have shown how the proton gradients Mitchell proposed are, indeed, causally relevant (Harold 1986; Cramer and Knaff 1990; Weber 1991; Allchin 1992a, 2996). The uncouplers that largely prompted Slater's original hypothesis, for example, are now viewed as eliminating the chemiosmotic gradient. Ernster's demonstration of reversed electron flow, too, is adequately explained through Mitchell's model: an electrochemical gradient assumes the role of the intermediate I~X (Figures 1, 2). Yet these conclusions do not wholly invalidate the causal relevance of observations once linked to the claims about the high-energy intermediates.

For example, the exchange reactions—once indirect evidence for the chemical hypothesis—are not fully addressed by chemiosmotic concepts. The ATP-ADP exchange reaction, once taken to reveal information about yet unidentified molecules, is now interpreted in terms of the enzyme that forms ATP. The current model of how the enzyme works is based on four exchange reactions, each used by the biochemists
in their earlier causal reasoning about intermediates. There are intermediates in this enzyme, but they are only found embedded in the reaction site, not independently in solution. They do not have high-energy bonds. The role of the intermediates in the overall process has thus been circumscribed, clarified in scope, and largely diminished. The intermediates exist, but in a much more limited domain. Attention to domain or scope, though, sensitizes one to see this aspect of the outcome, easily overshadowed by the dominant chemiosmotic theory (Figure 5b).

Finally, what became of Lehninger’s reasoning about the organization of the electron transport chain, part of how biochemists responded to criticism about the role of the membrane? This issue is still unresolved. The positions of the proteins in the membrane have been determined more exactly, but their functional arrangement is still not fully understood. The billiard-ball model based on how other enzymes interact, in this case, does highlight some important causal elements in the domain of ox-phos that the chemiosmotic theory does not.

Since its formal introduction in 1953, the chemical hypothesis has certainly dwindled in scope. The role of the high-energy intermediates, once central, has vanished. In addition, the shift to the chemiosmotic framework has entailed--as depicted in Kuhn’s (1962) model--a radical conceptual and experimental gestalt switch from earlier biochemical practice in ox-phos. Yet Mitchell’s alternative theory and experimental gestalt did not wholly eclipse all aspects of the chemical hypothesis or its domain. Many of the findings that initially led biochemists to search for the intermediates were "composted" into other areas of research practice or domains that are not addressed by the chemiosmotic model. Knowledge of the causal mechanism of the ATP enzyme, for example, depends in part on the results of the exchange reactions. The domain of ox-phos, once construed to be cohesive and interpretable within one theory or conceptual scheme (Figure 5a), unexpectedly became divided. The once comprehensive chemical hypothesis became fragmented across different, more isolated domains, while the chemiosmotic hypothesis displaced it in the "central" area of the ox-phos domain. Several complementary models, each mapping a relatively isolated domain, now explain collectively how energy is processed in the cell (see Figure 5b).

From the perspective of the chemiosmotic paradigm, the chemical hypothesis as a whole now seems "as relevant as phlogiston." When viewed more closely, however, its many "errors" fit into several small domains that are still part of interpreting and investigating ox-phos. More importantly, perhaps, the distribution of "error" allows one to see the overall pattern of the resolution of disagreement more clearly. The controversy was resolved because researchers differentiated the overlapping domains (Figure 5b), not because one theory replaced or outcompeted the other (as a result of, say, weight of evidence or some "crucial experiment"; see also Allchin 1994b). Ultimately, the Ox-Phos Controversy, like the fate of many proposed high-energy intermediates, was more about domains than hypotheses.

10. Reassessing Phlogiston

The differentiation of domains in the case of the "20th-century phlogiston," along with
the residual role of the chemical hypothesis, may seem at odds with its 18th-century counterpart. The Chemical Revolution is widely portrayed as "the overthrow of the phlogiston theory" (e.g., Conant 1957; Cohen 1985, p. 231; Donovan 1988). Phlogiston is supposedly an artifact, part of a deeply entrenched misconception of why things burn (e.g., Thagard 1990, esp. pp. 184, 201). The analogy with ox-phos may thus seem ill cast. However, fruitful analogies can also work in reverse (see note 1). Here, the 20th-century example can indeed serve as a valuable tool for reassessing the more familiar 18th-century debate.

Historically, the concept of phlogiston was ultimately abandoned. Careful attention to domains, though, can sensitize one to several late defenses of phlogiston—after and in full recognition of the discovery of oxygen. When one considers the domain including chemical composition and balanced weights in chemical reactions (and elemental nomenclature) as primary, phlogiston is indeed irrelevant or artifactual. In the domain including oxygen and Lavoisier's "doctrine of gases," phlogiston is likewise largely uninformative. However, within the domain of what we now call oxidation-reduction reactions, reduction potential, and energy, explanations using oxygen, etc., are insufficient (see also Brown 1864). Of course, phlogiston originally organized causal thinking in just this domain: unifying combustion, reduction, calcination (and later photosynthesis). The transfer or release of phlogiston offers a very simple causal model for oxidation and reduction reactions—-one that, even today, introductory chemistry students can appreciate and find compatible with explanations using oxygen (Allchin, 1997).

One should not be surprised to find, therefore, that late phlogistonists focused on precisely these elements (Allchin 1992b; 1994a). They discussed heat, light, phosphorescence, animal heat, coal, electricity, and their causal relationships to chemical reactions. At the same time, they scorned Lavoisier's preoccupation with weight—clearly implying that combustion did not fall exclusively within its domain. Though the concept of phlogiston did not survive this period historically, it did address a domain of energy interactions that most biochemists did not continue to pursue. One may well contend that, in this sense, the Chemical Revolution was less a replacement of the concept of phlogiston by oxygen, and more a dramatic displacement, based on exploring the new domain of weight relations opened by the discovery of oxygen and setting aside questions of energy highlighted by phlogiston. As many late phlogistonists argued (in other terms), one needed to differentiate the two domains.

One may also note that, like their 20th-century cousins, 18th-century phlogistonists could manipulate their entity, or 'principle', in the laboratory. That is, they could "intervene" (Hacking 1983). For example, they could exhaust the supply of phlogiston in a metal by heating it and forming its earthy calx. They could then reintroduce phlogiston from another phlogiston-rich source, such as charcoal, and reform the metal from the calx. Using an acid on a metal, they could trap the escaping phlogiston in a gaseous form and burn it. They could even use the phlogiston from electricity to reduce metals—-an effect that they earlier predicted based on the similarities between burning and electrical discharge (Sudduth 1978). Like ox-phos biochemists, phlogistonists had an effective scheme for intervening in nature. In this domain, at least—-underscored by late phlogistonists--phlogiston was far from being an
artifact.

One may interpret the success of intervening with phlogiston in two ways. First, one may view it as validating the claims for phlogiston: phlogiston was--and perhaps still is--"real." If so, then phlogiston "exists" only in the context of a limited domain where it can guide simple interactions--not in the domain of oxygen, etc. (This was how late phlogistonists argued). Second, one may regard phlogiston, like the high-energy intermediates of ox-phos, as an example of error (Lavoisier's posture.) Adopting this position, however, means that Hacking's popular argument fails: intervention itself is not a sufficient gauge for assessing the ultimate reliability or relevance of knowledge claims.

11. Conclusion

In both the 18th- and 20th-century episodes described above, domains items were dramatically regrouped and partly redescribed (§§9-10). Newly shaped domains also emerged to prominence, and with them, new problems. In both cases, the depth of the reorganization and its experimental consequences were revolutionary. The cases suggest more generally how analysis of domains in episodes of conceptual change and controversy may be fruitful. What may appear or be cast by scientists as a simple case of the linear succession of theories or "either-or" debate may, when one considers domains more closely, involve both "competing" theories as part of the outcome. Disagreement about conflicting hypotheses may be resolved by differentiating their domains.

Both cases have gained notoriety as examples of error. The history of the Ox-Phos Controversy offers clues to interpreting how that error was actively constructed (§3-5). From certain perspectives, Pinchot's, Sanadi's, Boyer's and others' claims about the high-energy intermediates (like phlogiston, too) became transformed into irrelevant artifact or error (§6). However, their findings became artifacts by being reconstructed into other facts (§7). One may specify the substantially different domains in which they are facts. From these other perspectives, the experimental methods and observations retain their relevance. Thus, one might say that error is in the eye of the beholder's domain. In underscoring the significance of domain in interpreting error and resolving debate, the analysis of "a 20th-century phlogiston" suggests the potential of further studies on the epistemology of error and on epistemic strategies for resolving disagreement.

Postscript: Methodological Overtones

The cases here provide a basis for broader methodological generalizations. First, being able to characterize errors (or artifacts) as themselves facts is an interpretive tool of widespread significance. Such an account is symmetrical, treating both claims in the same explanatory framework. It thereby addresses the challenge of the Strong Programme (Bloor 1976, p. 5). Both "true" and "false" claims may be facts, each situated in a specific domain. At the same time, the account is also empirical. Discussions of domain focus on observations and experimental relevance. There is a
context to each claim, but the context is not social, cultural, institutional, or even theoretical. "False" claims, as well as "true" ones, may be explained empirically—an alternative to regarding the principle of symmetry exclusively in sociological terms. The concept of domain allows one to follow a principle of symmetry meaningfully in empirical terms.

Second, a symmetrical approach involving domains accommodates the philosopher's aim of interpreting justification. The primary task of justification shifts, however, from characterizing evidence to characterizing the domain of a suite of evidence. The central question is no longer whether or how concepts are justified. Rather, one seeks to characterize the domain or scope within which a causal claim can fit or be deemed acceptable. Through this orientation, experiments need not merely "test" a theory, model or schema, or create a phenomenon. Instead, experiments may function to help articulate domain or relevant parameters (Shapere 1984, p. 284). One may appreciate that a substantial part of experimental work is establishing the domain or scope within which a concept may be justified—and not merely as a part of 'normal science' (Kuhn 1962, pp. 24-25, 29-30, 42). The experimental strategies that scientists use to reason about domains form a topic ripe for further study.

Though symmetrical accounts of process may be fundamental, scientists also create asymmetries. There were (as noted) important asymmetries in ox-phos and late 18th-century chemistry. The sixteen claims for the high-energy intermediates were not accepted. The chemical hypothesis did not continue to provide the central framework for studying ox-phos. Phlogiston was not considered part of explaining changes in the air or weight during combustion. The asymmetry in each case was positional or perspectival, however, not epistemological. As detailed above, claims of fact and error reflect a position and a primary domain. We can thus begin to discern exactly how the empirical position, or standpoint, is critical to interpreting claims of scientific knowledge (c.f., Harding 1991). 'Empirical perspective' or 'empirical context' are terms that can have meaning.

A related aim for someone using a principle of empirical symmetry, then, is to understand how each scientist, in social or cognitive terms, comes to occupy a particular context described in empirical terms. Empirical and cognitive (or social-biographical) accounts are not exclusive alternatives. Rather, they intersect and necessarily complement one another (Giere 1988; Nersessian 1991; a sample analysis for the phlogiston case may be found in Allchin 1994a; 1996). The principle of empirical symmetry can thereby bridge two major perspectives on science.

Third, the concepts of domain and empirical symmetry have implications for philosophical models of theory-choice. Models that are framed in either-or, winner-take-all terms are not effective for interpreting how the scientists resolved the Ox-Phos Controversy (or the Chemical Revolution, fully considered). A more sophisticated model, suggested by these episodes, includes the possibility of differentiating domains and redistributing sets of data (see Allchin 1994b; 1996). Scientists do disagree. In a case such as the ox-phos episode, an unqualified pluralism would be an inappropriate guide. We cannot escape the need to choose between incompatible and conflicting schemata or models. Rather than ask, "how do we choose between given alternatives?", one may ask more simply, "how do the interpretations conflict?" How
does one reconcile them with each other and with the data?” That is, problems about theory-choice may be conveniently reframed as questions about resolving disagreement. Reconciliation or accommodation of alternative views, rather than exclusive (“either-or”) theory-choice, is the aim.

Fourth, historiographically, the case of the “20th-century phlogiston” suggests the limits of exclusively linear narrative. A complete accounting of the ox-phos episode, for example, involves both how the chemical and chemiosmotic lineages converged, and ultimately how they diverged, tracking separate domains. The Ox-Phos Controversy and the disagreement over the high-energy intermediates is just one focal point for seeing how various lineages interacted. The case thus supports a more reticulated view of history. In this view, narratives adopt a richer pattern of multiple beginnings and multiple endings, with parallel stories knotted together on occasions by their various interactions.

Finally, the interpretive strategy that shaped the story of the ”20th-century phlogiston” has broader applicability. One can construct two narratives of the high-energy intermediates of ox-phos. The first is decidedly Whiggish and retraces their history from their present status as a noteworthy case in error. The second account emerges from the perspective of their original status as a fact or potential fact. The challenge is for the historian to maintain historical sensitivity to this particular perspective and to exercise historical imagination while interpreting events that later unfolded. When one traces the history forwards from the now discredited fact, one perceives how the data for phosphohistidine, NAD-E, etc., do reflect facts, though in domains not obvious from a retrospective view. These facts are peripheral and irrelevant given the first perspective, and thus go unnoticed. The second history is just as "biased" or shaped by a certain perspective as a Whiggish account. But the guiding viewpoint or "lens" is situated historically and looks forward, not backward. One might therefore call this 'reverse Whiggism'. For example, from whose perspective do we view the Chemical Revolution? How did the outcome appear to late phlogistonists, for example? Understanding the "rear guard" in a scientific revolution honors a principle of symmetry and a historian's aim in interpreting claims in their historical context. But by tenaciously following the rear guard as debate resolves itself--refusing to succumb to the current view--one can highlight different, possibly unappreciated elements of the story. Reverse Whiggism may thereby offer a widely productive tool for historical discovery.

Appendix: Interpreting Domains

A domain, conceived from the perspective of scientific practice and discourse, is the constellation of related phenomena--such as might be the focus of an investigation, a series of investigations, or, more broadly, a research lineage--that scientists often strive to manipulate, interpret causally, model, or explain. In the metaphor where theories are "maps" (e.g., Turnbull 1989), a domain is the "territory." One may fruitfully characterize the domain concept in other specific contexts, as well, reflecting the interconnectedness of different dimensions of or perspectives on science.

Experimentally, a domain will reflect all the techniques or processes that need to be enlisted to produce particular phenomena or to "construct" particular observations.
These means of accessing a domain appear prominently in a scientific paper or research report in the Methods section. The variables that need to be controlled (in the sense of excluded), or those that are considered irrelevant, are also addressed in these sections. The frequent care in excluding certain elements reflects the importance of identifying elements as inside versus outside a domain. The broader experimental context in which the constellation of variables or methods is set--typically described in the Introduction of a paper, and frequently addressed again in the Discussion or Conclusions--can be equally important. In these sections, researchers nest the local domain of their experiments in the broader domain of other experiments and observations, relying on numerous conventional categories, concepts, and theories in the process. Knorr-Cetina (1981, pp.94-135), Latour (1987, pp. 50-52) and Myers (1990, pp. 63-100), among others, have emphasized the interpretive aspect of this enterprise and the potential disagreement possible in specifying domains of broader scope.

In a stronger perspective, one may want to define a domain instrumentally. Ackermann (1985) follows this strategy to great effect through his notion of 'data domains'. This approach is valuable in underscoring the means of observations, rather than the observations themselves (or interpretations of them). The corresponding problem with delineating domains solely in terms of certain apparatus or techniques, of course, is that new apparatus and techniques necessarily create new domain, even when we might construe them as merely providing alternative access to the "same" phenomenon. It becomes difficult, for example, to relate several different assays for measuring the amount of ATP produced by the same series of chemical reactions (§6). The limits of this instrumental view, of course, parallel the well documented problems of 'operational' definitions for logical empiricists.

Philosophically, the most important dimensions of a domain may be explanatory. What is it that a scientist intends to explain? Why are some elements included, while others are omitted in explanations? Here, the domain includes all that is explanatorily relevant. To the extent that an explanation is evidentially grounded, one will undoubtedly encounter tensions between conceptual and empirical dimensions. Presumably, though, each concept "maps" onto a suite of specific elements of observational domain. A philosopher's view of domain, of course, may differ from the experimentalist. A philosopher is likely sensitive to conceptual dimensions or contexts that remain opaque (or are simply assumed) in a community of experimenters. A philosopher may well characterize the domain with broader borders than an individual researcher might, while still referring to essentially the same domain.

The concept of domain may also be interpreted socially. What brings scientists into communities and holds them together in working relationships? One fundamental element is shared domain. Social relationships and exchanges emerge between individuals--even those who are loathe to develop any congenial mutualism--because they deal with the same set of phenomena in their investigations. The key relationships of who reads whose papers, who cites whose papers, who gets samples from whom, who goes to whose lab to learn certain techniques, who goes to the same scientific meetings, etc., all build on the need to communicate or share results (both conceptually and materially). Thus, causal domains--built experimentally--may well be further
expressed sociologically. While I have not emphasized this dimension in my analysis, the Ox-Phos Controversy was social as well as conceptual. Afterwards, three groups of researchers each retreated into their relatively isolated professional circles: one group examined the chemistry of cytochromes, for example; a second group studied enzymological mechanisms of the ATP enzyme; and a third group largely investigated membrane-related phenomena (see §9). Still, while they investigate the mechanisms of particular aspects of the process, they need to work in full awareness of their "neighbors." In addition, other researchers from previously distinct fields now cluster together to solve new sets of problems in the domain structured conceptually by chemiosmosicory theory (see Figure 5b).

References


Chance, Britton and Brigitte Schoener. 1966. "High and Low Energy States of
Mechanism in Mitochondria: Kinetic, Spectroscopic and Thermodynamic
University Press (Belknap).
Conant, James Bryant. 1957. "The Overthrow of the Phlogiston Theory." In Harvard
University Press.
the Intermediates in a Model Oxidative Phosphorylation Reaction." J. Biol
Cooper, Cecil and Albert L. Lehninger. 1957. "Oxidative Phosphorylation by an
Enzyme Complex from Extracts of Mitochondria. IV. Adenosine Triphosphatase
Activity. V. The Adenosine Triphosphate-phosphate Exchange Reaction." J.
Biol. Chem. 224:547-78.
Cramer, William A. and David B. Knaff. 1990. Energy Transduction in Biological
the Coupling Sites of Mitochondrial Oxidative Phosphorylation." Biochem.
Phosphorylated Intermediate in Mitochondrial Oxidative Phosphorylation."
Culp, Nancy. 1994. "Defending Robustness: The Bacterial Mesosome as a Test
Case." Pp. 46-57 in PSA 1994, vol. 1. Edited by David Hull, Micky Forbes and
Osiris, vol. 4.
Baltimore: Johns Hopkins University Press.
Ernster, Lars. 1963. "The Phosphorylation in the Flavoprotein Region of the
Macmillan (Pergamon).
Biol. 91:227S-255S.
Franklin, Allan. 1990. Experiment, Right or Wrong. Cambridge University Press.


