Synthese (2012) 185:125–144
DOI 10.1007/s11229-011-9879-7

# Life without definitions

**Carol E. Cleland**

Received: 24 June 2010 / Accepted: 6 January 2011 / Published online: 15 February 2011

**Abstract** The question 'what is life?' has long been a source of philosophical debate and in recent years has taken on increasing scientific importance. The most popular approach among both philosophers and scientists for answering this question is to provide a "definition" of life. In this article I explore a variety of different definitional approaches, both traditional and non-traditional, that have been used to "define" life. I argue that all of them are deeply flawed. It is my contention that a scientifically compelling understanding of the nature of life presupposes an empirically adequate scientific theory (vs. definition) of life; as I argue, scientific theories are not the sort of thing that can be encapsulated in definitions. Unfortunately, as I also discuss, scientists are currently in no position to formulate even a tentative version of such a theory. Recent discoveries in biology and biochemistry have revealed that familiar Earth life represents a single example that may not be representative of life. If this is the case, life on Earth today provides an empirically inadequate foundation for theorizing about life considered generally. I sketch a strategy for procuring the needed additional examples of life without the guidance of a definition or theory of life, and close with an application to NASA's fledgling search for extraterrestrial life.

**Keywords** Life · Definition · Theory · Theoretical identity · Natural kinds · Anomalies

## 1 Introduction

In earlier work (Cleland and Chyba 2002, 2007; Cleland 2006), directed primarily at an audience of scientists, I argued that the popular definitional approach to explaining

---

C. E. Cleland (✉)
Department of Philosophy, Center for Astrobiology, University of Colorado, Boulder, CO, USA
e-mail: Cleland@colorado.edu

the nature of life is fundamentally misguided. This article refines and further develops these arguments within the context of contemporary developments in philosophy of language and philosophy of science. It is divided into two parts. In Sect. 2, I explore a variety of traditional and non-traditional definitional approaches to understanding life. I argue that they are all deeply problematic from both a philosophical and a scientific perspective. In Sect. 3, I explore an alternative, scientifically more promising approach to explaining the nature of life and sketch some of its consequences for NASA's fledgling search for extraterrestrial life.

I begin in Sect. 2.1 by discussing the scientific implausibility of a definitional approach to explicating the nature of life. Because definitions, in the traditional logical sense, are primarily concerned with language and concepts they do not provide good tools for answering scientific queries about natural kinds, i.e., categories that are in some sense carved out by nature, as opposed to human interests and concerns. Even more problematically for both traditional and non-traditional definitions of life, our current scientific understanding of life is too limited to support a theoretically useful, tentative generalization about all life.

As I discuss in Sect. 2.2, the little support that can be garnered for the traditional project of defining 'life' derives from the descriptive theory of the meaning and reference of natural kind terms. In Sect. 2.2.1, I explain why philosophers of language no longer accept this theory. Section 2.2.2 explains why none of the popular alternatives to the descriptive theory (viz., cluster theories, causal theories, and hybrid descriptive-causal theories) supports a definitional approach to understanding natural kinds. I conclude that if (as seems likely but not certain) life is a natural kind, the traditional definitional approach to understanding the nature of life is unlikely to be successful.

Some philosophers and many scientists employ the term 'definition' in a nontraditional way to include empirically revisable, provisional theoretical claims about life. These "theoretical definitions" are sometimes characterized as theories. As I discuss in Sects. 2.1 and 3.1, however, theoretical definitions share many of the defects of traditional definitions and few of the advantages of *bona fide* scientific theories, which are much more holistic than definitions. I conclude that neither a traditional nor a non-traditional definitional approach to answering the question 'what is life?' is likely to be successful. I argue in Sect. 3.1 that an answer to this scientifically important question presupposes an empirically adequate, authentic scientific theory of life.

As discussed in Sect. 2.1, the philosophical difficulties faced by the various definitional approaches pale in the face of an even more serious empirical problem facing anyone wishing to generalize (whether by means of a definition or theory) from known life to all life. Biochemists, molecular biologists, and microbiologists have discovered not only that familiar Earth life represents a single example, in virtue of sharing a last universal common ancestor, but also that it may well be unrepresentative of life in general, reflecting chemical and physical contingencies present on the early Earth. One cannot safely generalize to all of life, wherever and whenever it may be found, from a single, potentially unrepresentative, example of life. As a consequence, scientists and philosophers are currently in no position to formulate even a tentative general theory of life. Indeed, as I discuss, endorsing a generalization about the nature of life prematurely is more likely to hinder rather than help the search for additional examples of life, blinding scientists to forms of life differing significantly from familiar

Earth life should they encounter them.[1] And yet the latter provide the best grist for the theoretical mill. In Sect. 3.2, I sketch a strategy for searching for unfamiliar forms of life without the aid of a tentative theory or definition of life. As I argue, this approach is more compatible with the way in which successful scientific theories have historically developed in light of empirical evidence.

## 2 Definitional approaches to understanding the nature of life

### 2.1 The scientific implausibility of definitional approaches

In the traditional sense, definitions are concerned with language and concepts. The definition "'bachelor' means unmarried, human male" provides the well-worn example from informal logic classes. This definition explicates the meaning of a word ('bachelor') by dissecting the concept associated with it by competent speakers of English into sub-concepts supplying (within the inevitable constraints of vagueness)[2] necessary and sufficient conditions for its application; in essence, the definition fixes the extension of 'bachelor'. When a natural scientist seeks an answer to the query "what is life?," however, she is not interested in an analysis of the contemporary human *concept* of life. She wants to know what life *truly* is: what bacteria, slime molds, fungi, fish, insects, trees, birds, and elephants all have in common that distinguishes them from nonliving physical systems, and what *any* system, however different, must have in common with these organisms in order to qualify as "life." Analysis of the twenty-first-century human concept of life is unlikely to be of much help in this endeavor *if* life is a natural kind, that is, a category that is in some (modest or robust) sense carved out by nature and would exist even if there had been no human beings to think or talk about it. Granting for the sake of argument that life is a natural kind,[3] if our contemporary

---

[1] As discussed in (Cleland and Copley 2005; Davies and Lineweaver 2005; Cleland 2007) it is possible that there are alternative forms of microbial life right here on Earth (a "shadow biosphere") that have not been recognized for what they represent because they do not resemble familiar Earth life closely enough to be detected by the tools currently used to explore the microbial world.

[2] Vagueness is now understood to be a pervasive feature of natural languages, frustrating even the most promising efforts to fully specify necessary and sufficient conditions for the application of a general term. Our definition of 'bachelor', which provides the prototype for full definitions, is no exception. It confronts borderline cases such as unmarried adolescent males or the Pope, and efforts to exclude these cases typically produce new borderline cases. What a definition needs to get right are the unambiguous cases, e.g., thirty-year old married and unmarried men. The moral is that some borderline cases are to be expected and hence do not do not pose a serious threat to definitions. It is thus a mistake to think that a satisfactory definition of 'life' needs to handle cases (e.g., viruses) that truly seem to straddle the border between life and nonlife. As we shall see, however, far more serious difficulties confront attempts to "define" life.

[3] While there are compelling (but not conclusive) arguments for believing that species and higher taxa might not be natural kinds, these arguments do not carry over *mutatis mutandis* to the distinction between living and nonliving physical systems. It is of course possible that life is not a natural kind or, for that matter, that anti-realists are right and there is no principled distinction between natural and non-natural kinds. It is beyond the scope of this article to engage in this debate, however. It suffices to say that most biologists believe that life is a natural kind; they believe that they are exploring a natural phenomenon that would exist even if *Homo sapiens* had not evolved on Earth. In this article, I assume that they are right about this, and take a realist stance. But as the reader will see, my realism need not be overly strong to make the case that a definitional approach to understanding life is fundamentally mistaken.

beliefs about life are incorrect, a definition of ‘life’ will serve only to entrench our misconceptions, making it more (not less) difficult to achieve a satisfactory scientific understanding of the nature of life.

Admittedly, not everyone who advances a “definition” of life has in mind the traditional notion of definition. This is especially true of scientists, who typically have little formal training in logic or philosophy. But it is also true of some philosophers. Mark Bedau (1998), for instance, articulates his “theory of life” in the form of a “definition” of life (p. 128). Such definitions are sometimes called “theoretical definitions” in order to distinguish them from traditional definitions. Theoretical definitions are characterized as tentative and revisable in light of empirical evidence. The problem with theoretical definitions of life is that they closely resemble traditional definitions in specifying necessary and sufficient conditions for life based upon current scientific beliefs about life.[4] As a consequence they face many of the same problems as traditional definitions. Because they fix (albeit tentatively) what sorts of things qualify as life and what sorts of things do not, they are just as likely to seriously mislead us as traditional definitions of ‘life’ if (as I shall argue is highly likely) our current scientific concept of life is unreliable. If one utilizes such a definition to guide attempts to “create” life in the laboratory or search for life on other worlds one is likely to produce or find only what one is looking for. Furthermore, interpreting theoretical definitions as “theoretical identity statements” couched in terms of (as opposed to encapsulating) scientific theories does not alleviate this difficulty. For as discussed below, we just don’t know enough about life as a general phenomenon to formulate scientifically fruitful identity statements about it. (As discussed in Sect. 3.1, interpreting theoretical definitions as theoretical identity statements faces additional difficulties as well.)

To fully appreciate the central difficulty facing both traditional and nontraditional definitional approaches to understanding life consider an analogous, although somewhat contrived, example from the history of science. Suppose that a seventeenth-century scientist attempted to answer the question ‘what is water?’ by “defining” water. This is before the advent of molecular theory, whose foundations were laid in the late eighteenth century by Antoine Lavoisier (1783). As a consequence, our hypothetical scientist knows nothing of molecules. His knowledge of water is limited to sensible properties such as being wet, transparent, tasteless, odorless, and a good solvent. Unfortunately many of the substances called “water” by his contemporaries lack one

---

[4] This isn’t the only way in which the term “theoretical definition” is used in contemporary philosophical discourse. Under the semantic conception of scientific theories, a “theoretical definition” defines a scientific theory as a whole (qua conjunction of axiom-like statements). In an informal discussion, Mark Bedau informs me that this is what he really means by a “definition of life.” The problem is that the theoretical definitions of the semantic conception do not define natural kind terms (such as ‘water’ or ‘life’) separately. They define whole theories subsuming a variety of different natural kinds, and it isn’t at all obvious that the natural kinds falling under a particular theory can be “defined” (supplied with necessary and sufficient conditions) individually within the framework of the theory. I discuss these issues in detail in my forthcoming book, *The Quest for a Universal Theory of Life: Searching for Life as we don’t know it.* For purposes of this paper, however, the point is that Bedau’s claim that his definition of *life* is a theoretical definition in the sense of the semantic conception is extremely problematic. In any case, when I use the term ‘theoretical definition’ in this paper I have in mind the more traditional sense of a definition supplying necessary and sufficient conditions for the application of a natural kind term within the context of the theoretical framework provided by a scientific theory; such definitions are also sometimes called ‘special’, ‘technical’, or ‘scientific’.

or more of these properties. Muddy water is not transparent, salty water is not tasteless, and brackish water is not odorless.

It is difficult for us, steeped in twenty-first-century chemistry, to fully appreciate the dilemma that our seventeenth-century scientist faces in selecting one or more of the sensible properties associated with water as essential to it. The alchemists, who were medieval chemists, were impressed by water's powers as a solvent, and hence chose solvency. As a consequence they classified nitric acid and mixtures of hydrochloric acid, which are even better solvents than water (and share some of its other sensible properties, e.g., being transparent and a liquid), as "water;" it is not an accident that the former was called "*aqua fortis*" (strong water) and the latter "*aqua regia*" (royal water) (Roberts 1994). But as we now know the alchemists were wrong. Nitric acid is not water, and like muddy water and salty water, mixtures of hydrochloric acid are not pure water. What distinguishes water from all other chemical substances is a unique molecular composition, $H_2O$. Being $H_2O$ is what salty water, muddy water, brackish water, distilled water, and even acidic solutions have in common, despite their phenomenal differences. It is what distinguishes nitric acid, whose molecular composition is $HNO_3$, from water despite their phenomenal similarities. Could a seventeenth-century scientist have discovered this? The answer is clearly "no." His scientific understanding of water is based upon its superficial sensible properties, and no amount of reflection on or investigation of these properties with the primitive equipment then available could reveal that water consists of two atoms of hydrogen chemically bonded to an atom of oxygen. To achieve the latter understanding of water he needs a new theoretical framework for thinking about chemical substances: molecular theory.

In some ways the situation is worse for contemporary scientists seeking an understanding of the nature of life than for our seventeenth-century scientist seeking an understanding of the nature of water. For there are compelling scientific reasons for thinking that life as we know it on Earth today represents a single, potentially unrepresentative, example of life. Molecular biologists have discovered that all known life on Earth descends from a last universal common ancestor (LUCA), most likely a community of primitive proto-cells, subject to rampant mutation, exchanging cellular components in a manner reminiscent of modern lateral gene transfer but unable to replicate with the fidelity required by natural selection (Woese 2004). This means that we are dealing with a single example of life. Without additional examples of life, one cannot discriminate features that are universal to life, wherever it might be found, from features deriving from mere physical and chemical contingencies on the early Earth, or representing mere quirks of the lucky little bugs that happened to make it across Woese's "Darwinian threshold" and become the ancestors of us all.

In addition, biochemists have established that some of the basic molecular building blocks of familiar life could have been at least modestly different without compromising their biological functionality; for more detail, see Benner (1994) and Benner et al. (2004). Life on Earth could have utilized a different collection of amino acids, for instance, to synthesize its proteins, which comprise the bulk of its structural and enzymatic material; familiar life employs the same approximately 20 amino acids, out of over 100 possibilities, to build its proteins. Similarly, life could have synthesized nucleic acids (DNA and RNA), which store and translate its hereditary material, out of a different suite of nucleotide bases or used different sugars to construct the sugar-phosphate backbones of nucleic acids.

The molecular similarities among extant Earth organisms do not end with their basic chemical building blocks. All known life on Earth utilizes a triplet coding scheme to store hereditary information on nucleic acids and, moreover, the same triplet of bases (known as a codon) always codes for the same amino acids; there is some redundancy with most amino acids being represented by more than one codon but the same codon never represents more than one amino acid. There is little chemical reason to suppose that life couldn't have mapped codons to different amino acids or utilized a different number of nucleotide bases (two or four, for instance, instead of three) to encode amino acids, particularly if it had employed a different number of amino acids to build its proteins. Furthermore, the process of synthesizing proteins from the information encoded on DNA is carried out in ribosomes, which are minuscule but highly complex molecular machines composed of both RNA and protein. It is unlikely that the earliest forms of life on Earth utilized something this sophisticated to build its structural and enzymatic material, and even more unlikely that ribosomes represent the only possible chemical mechanisms for performing this crucial biological function. These remarkable molecular similarities among all extant Earth organisms most likely reflect contingent characteristics of LUCA, their last universal common ancestor.

Clearly, familiar Earth life must have something in common with all life if (as I am assuming for the purposes of this essay) life is a natural kind. The problem is that our scientific ideas about life are based upon a single example that may not be very representative of life. This is true not only of the molecular characteristics of familiar life but also of its abstract functional characteristics, two of which are commonly held up as providing the most promising candidates for essential properties of life. These two characteristics are (i) the capacity to self-organize and maintain self-organization for an extended period of time against internal and external perturbations and (ii) the capacity to reproduce and transmit to progeny adaptive heritable modifications. The notion that these characteristics provide the best candidates for essential properties of life rests upon the assumption that life is a functional as opposed to compositional or structural kind. But there is little empirical evidence to support the assumption that life is a functional kind. For all we know these pervasive functional characteristics of contemporary Earth life represent unreliable symptoms of more fundamental but as yet unknown properties of life. This sort of thing has happened before in science. Some diseases (e.g., tuberculosis and bubonic plague) identified as distinct before the advent of the germ theory of disease on the basis of symptoms are now recognized to be the same in virtue of being caused by the same microorganisms; some microorganisms produce strikingly different symptoms when they infect different organs (e.g., lungs vs. lymph glands). Similarly, as discussed earlier, different chemical substances (qua distinct molecular compounds) sharing some of the same sensible properties (e.g., being liquid, transparent, and a good solvent) were once mistakenly classified as the same chemical substance. Thus although Ruiz-Mirazo and Colleagues (2002) are correct in pointing out that we possess an enormous amount of empirical information about life they are wrong in inferring that this body of information is adequate for generalizing beyond familiar Earth life.

In summary, from a scientific perspective, explicating the nature of life by means of either a traditional or nontraditional definition of life is fundamentally misguided. Our experience with life is limited to a single example that we have good scientific reasons for believing could have been at least modestly different. Moreover we have no idea how different life could be from life as we know it on Earth today. Until we encounter forms of life descended from a separate origin we will be in no position to speculate about the possibilities for life considered generally. (This might be thought to give rise to a quandary: How can scientists search for alternative forms of life without the guidance of a definition or theory of life? I return to this issue in Sect. 3.2.)

## 2.2 The philosophical implausibility of definitional approaches

As we have seen, there is little scientific support for a definitional approach to understanding natural kinds such as water and (presumably) life. During the first half of the twentieth century many philosophers of science nevertheless interpreted claims such as ‘water is $H_2O$’, ‘sound is a compression wave’, and ‘heat is a motion of molecules’ as stipulative (a.k.a. “special scientific”) definitions. As traditionally understood, stipulative definitions represent arbitrary decisions to assign new concepts to old or new terms. It is clear that the meanings of some scientific terms are acquired in this fashion. Coining the term ‘gene’ for the basic unit of heredity provides a good example of assigning a new term to a theoretically important concept in biology. The physicist’s definition of ‘work’ as the product of the magnitude of an acting force and the displacement due to its action, on the other hand, provides a good example of assigning an old familiar term to a theoretically important concept in physics. It isn’t clear, however, that identity statements such as ‘water is $H_2O$’ are stipulative in this sense. The decision to use the common term ‘water’ for the molecular compound $H_2O$ was not arbitrary. Armed with a new theoretical framework, molecular theory, nineteenth-century chemists engaged in extensive empirical research to discover the molecular compositions of chemical substances such as water; indeed, it is conceivable that the outcome of their investigations was that water is not $H_2O$. The question is why would someone interpret theoretical identity statements such as ‘water is $H_2O$’ as stipulative definitions?

### *2.2.1 The descriptive theory*

The view that theoretical identity statements such as ‘water is $H_2O$’ amount to stipulative definitions derives from an old philosophical theory of general terms known as the descriptive theory. Commonly attributed to John Locke, the descriptive theory dominated philosophical thought about the meaning and reference of natural kind terms through the mid-twentieth century and, despite the fact that it is no longer widely accepted among philosophers of language, remains influential today. According to the descriptive theory, the meanings of all general terms (including natural kind terms) are fully exhausted by the (mental) concepts associated with them. On the classic version of the theory, concepts are identified with descriptions *qua* logical conjunctions of predicates supplying necessary and sufficient conditions for the application of the term. The concept associated with a general term thus determines what falls into its extension; anything satisfying the description falls into its extension and anything that does not satisfy the description does not fall into its extension. This rules out the possibility of discovering that one is wrong about the nature of the kind designated by a natural kind term. If one changes the concept (description) associated with a phenomenal kind term such as water one *ipso facto* changes its meaning and reference (extension). The upshot is that nineteenth-century chemists cannot be viewed as having discovered that Aristotle, who believed that water is an indivisible element, was wrong about phenomenal water. Yet surely this is exactly what happened.

In this context, it is important to keep in mind that embracing a theory of meaning and reference has serious ramifications for our understanding of how scientific theories (which are linguistic entities) connect the entities and processes that they postulate to the perceptual phenomena of ordinary, everyday experience that they are supposed to explain. If the descriptive theory is correct then theoretical identity statements cannot be interpreted as representing authentic discoveries about phenomenal kinds because the concept associated with the phenomenal kind term has changed and concepts *qua* descriptions fully exhaust the “nature” of the items falling into the extension of the term. In other words, on the descriptive theory, water doesn’t have a nature independently of the concepts that humans, including scientists, associate with the term ‘water’. Change the concept and you *ipso facto* change the “nature” of the kind. This helps to explain why so many early-twentieth-century philosophers of science felt compelled to do violence to their intuitions and interpret theoretical identity statements as stipulative definitions analogous to the physicist’s definition of ‘work’.

From its inception, the descriptive theory of general terms faced problems. It is unable to distinguish natural kind terms from non-natural or artificial kind terms such as ‘bachelor’, ‘garbage’, ‘hammer’, and ‘American’; unlike natural kind terms, these terms designate categories or classes of items that are carved out by human conventions, interests, and concerns, and which would not exist had there been no human beings. Locke was fully aware of this difficulty, and chose to bite the bullet (so to speak) and reject the distinction. In a revealing discussion (Locke 1689, BK. III, Chap. XI, Sect. 7), Locke argues that the seventeenth-century debate over whether bats are birds has little scientific merit, since the seventeenth-century *concepts* of bat and bird are compatible with either position; for Locke, the debate is merely verbal. In essence, Locke is demanding a stipulative definition of ‘bat’. In hindsight, however, this seems misguided. The question of whether bats are birds is not just a matter of deciding which descriptions to associate with ‘bat’. Scientists have discovered that the creatures called “bats” are more closely related to mammals than to birds. On the classic version of the descriptive theory, however, we must deny that this scientifically important claim represents a *bona fide* discovery.

Kuhn’s infamous argument for the incommensurability of rival scientific theories (1962, Chap. IX) is grounded in the classic version of the descriptive theory. As Kuhn points out, the descriptions associated with certain basic scientific terms undergo radical change during scientific revolutions. In Newtonian dynamics, for example, mass is conserved, whereas it is interconvertible with energy in Einstein’s special theory of relativity. To the extent that these “definitions”—Kuhn’s own word (p. 102)—of ‘mass’ are logically incompatible, ‘mass’ must be viewed as designating a different category of stuff in the two theories; ‘mass’ means something different in Einstein’s theory than it does in Newton’s theory. Kuhn concludes that one cannot (as physicists typically do) claim that Newtonian dynamics is a special case of Einsteinian dynamics. Instead of extending Newton’s theory to cases involving speeds close to that of light, Einstein has merely changed the subject (the class of items being talked about). But if (like most physicists) one views ‘mass’ realistically (vs. instrumentally), as designating a basic (albeit imperceptible) natural kind distinct from, for example, charge, then this seems wrong; Einstein taught us something new about mass that Newton didn’t know.[5] Analogously, on Kuhn’s view, nineteenth-century chemists cannot be construed as discovering something new about the same old stuff (water) that the alchemists or, for that matter, Aristotle were talking about; they merely changed the subject, redefining ‘water’. As with Locke’s account of bats and birds, these characterizations seem fundamentally mistaken. We do not think or speak about most scientific revolutions in this manner. We view them as representing defeasible discoveries, typically about categories of items (water, bats, gold, etc.) to which human beings have historically (albeit loosely) referred and that are (in at least some sense) delimited by nature as opposed to human beings. However unlikely it might now seem, we cannot completely rule out the possibility of future chemists discovering that water is not $H_2O$.

The success of a theory of meaning crucially depends upon its ability to account for indisputable facts about language and thought, for the subject matter of a theory of meaning is none other than language and thought. The failure of the classic version of the descriptive theory to account for the ways in which we speak and think about natural kinds is enough to seriously undermine it. But it is not catastrophic. Perhaps the differing ways in which we think and speak about natural kinds and non-natural kinds can be explained in a more holistic manner.

Efforts to patch up the classic version of the descriptive theory for the meanings of natural kind terms were defeated by Putnam (1975). Although it is beyond the scope of this article to explore his arguments in detail, it is important to appreciate the strength of Putnam’s case against the classic theory, particularly in light of the fact that a surprising number of philosophers of science still seem sympathetic to definitional approaches to understanding natural kinds. Part of the problem is that Putnam’s arguments against the classic theory are sometimes misconstrued as supporting his alternative, the causal theory of reference. While there is reason to doubt the viability of a (purely) causal theory of reference for the meanings of natural kind terms, there is little doubt that Putnam’s argument against the classic version of the descriptive theory succeeds.

Putnam encapsulated his central argument against the classic theory in the context of a now-famous thought experiment. He asked us to imagine a fantastic “planet,” Twin-Earth, which is just like Earth except that the stuff called “water” has a

---

[5] On an instrumentalist account of theoretical kinds Kuhn’s argument goes through for ‘mass’ (which was introduced as a theoretical term in Newtonian mechanics), but not for phenomenal kind terms such as ‘water’, because mass does not have a “nature” independent of what the theory says about it. The question of whether theoretical kinds should be given a realist or instrumentalist interpretation is unfortunately beyond the scope of this article; for an interesting discussion, however, see Psillos (1999, Chap. 12).

different chemical composition; he abbreviated the formula of the stuff on Twin-Earth as "XYZ." XYZ and $H_2O$ have the same sensible properties, namely, being wet, transparent, odorless, tasteless, and a good solvent. The oceans, lakes, and rivers of Twin-Earth contain XYZ, as opposed to $H_2O$. Twin-Earth is occupied by people who are not only biologically, psychologically and culturally like us but whose history also closely parallels our own.

The chemical composition of the stuff called "water" on Twin-Earth and Earth is not discovered on either planet until the early nineteenth century. Let $\text{Oscar}_E$ be a typical speaker of English on Earth in 1750 and let $\text{Oscar}_T$ be a typical speaker of English on Twin-Earth. $\text{Oscar}_T$ and $\text{Oscar}_E$ not only look alike but also have the same backgrounds, experience the same feelings, and (most importantly) have the same mental (psychological and neurophysiological) states. Because the stuff called "water" on both planets has the same sensible properties, Twin-Earthers and Earthlings have the same concept of water; they use the same identifying descriptions to pick out samples of "water" on their respective planets. Were $\text{Oscar}_E$ to visit $\text{Oscar}_T$, he would most likely believe that there is water on Twin-Earth. But as scientists on Earth and Twin-Earth discovered in the nineteenth century, he would be wrong. The stuff that $\text{Oscar}_T$ calls "water" has the wrong chemical composition. It is not $H_2O$. It follows that the extension of the term 'water' is not fully determined by concepts in the mind. If it were we would draw a different conclusion from the Twin-Earth example. We would conclude (on the basis of the sameness of their concepts) that $\text{Oscar}_E$ and $\text{Oscar}_T$ are talking about the same kind of stuff. Putnam has driven a wedge between the referents of natural kind terms and the concepts (*qua* identifying descriptions in our minds) that we associate with them. As he colorfully put it, "[c]ut the pie any way you like, 'meanings' just ain't in the head" (1975, p. 277).

It is important to be clear about what the Twin-Earth thought experiment does and does not establish. It establishes that the concepts associated with *some* natural kind terms do not completely determine their referents. The fact that it makes little scientific sense to speak of Twin-Earth as being just like Earth except for the chemical composition of water is irrelevant. Putnam is making a point about language and concepts. A theory of meaning should be neutral with respect to what our scientific theories tell us about the world. Language is used to describe many different kinds of situations, from factual, to hypothetical (e.g., what if Al Gore had been the U.S. President in 2003), to fantastic (e.g., the adventures of the young wizard Harry Potter). It is the responsibility of a theory of meaning to account for the ways in which we think and speak about fantastic situations as well as factual and hypothetical situations. Putnam's Twin-Earth thought experiment utilizes a fantastic situation to demonstrate an important point about language: Concepts (*qua* descriptions in our "heads") do not always fully determine the extensions of natural kind terms.

The Twin-Earth example has been extensively criticized. Although some, for instance, Mellor (1977) and Zemach (1976), have challenged Putnam's intuition that XYZ is not water, it is significant that their counter intuitions are not widely shared. The most that can be said is that we are uncertain about what to say in such a "fantastic situation" (Dupré 1993, p. 25). But, and this is the crucial point, this uncertainty alone is enough to undermine a definitional approach to understanding natural kinds.

If the classic theory were correct, we should not hesitate in drawing the conclusion that Mellor and Zemach urge upon us.

### *2.2.2 Alternatives to the classic version of the descriptive theory*

Most criticisms of the Twin-Earth example are directed at two additional theses that Putnam, along with Kripke (1972), advances: (1) a (purely) causal theory of reference and (2) a microstructural account of the nature of chemical substance, and by conversational implicature, all natural kinds. As discussed below, neither of these theses is secured by the Twin-Earth example.

At the heart of the causal theory of reference is the idea that the extension of a natural kind term is fixed not by human concepts but the actual nature of the things the term designates. As a consequence, a speaker need not fully understand the nature of a kind in order to succeed in referring to it. Someone who doesn't know that water is $H_2O$ can nonetheless succeed in referring to it, and nineteenth-century chemists can change their concept of water in light of new empirical and theoretical developments without ceasing to refer to the same old stuff that their predecessors were referring to when they used the common term 'water'. The central problem facing proponents of the causal theory is how to flesh out the requisite referential relation between bits of language and bits of the world. The referential relation is typically analyzed into two components. The first involves an initial reference-fixing event involving appropriate causal interactions with particular samples of the kind, either through perception or the use of criterial (vs. defining) properties extrapolated from expert sanctioned, stereotypical samples of the kind. Reference is subsequently transferred from speaker to speaker in virtue of their uses of the kind term lying on an appropriate causal chain stretching back to the initial reference-fixing event.

Despite their promise, causal theories of reference face serious difficulties that have yet to be overcome. Fortunately, the causal theory is not the only alternative to the classic version of the descriptive theory that is able to circumvent the Twin-Earth puzzle. The oldest alternative to the classic version of the descriptive theory is the cluster theory. Drawing its inspiration from Wittgenstein's famous analysis of the meaning of 'game', the cluster theory follows the classic theory in identifying concepts with descriptions, but it analyzes them as loose clusters (vs. logical conjunctions) of predicates. On the cluster theory, there are no sufficient conditions for belonging to a kind, and no single description can be taken as necessary. Being a member of the class of items designated by a natural kind term is just a matter of having "enough" (different versions tell different stories) of the pertinent properties. While not completely excluding the possibility that the concepts (*qua* descriptions) associated with some natural kind terms provide necessary and sufficient conditions for their applications, the cluster theory denies that this is typically the case, which is enough to frustrate the project of defining natural kind terms. In short, the cluster theory is a version of the descriptive theory that rejects a definitional approach to understanding natural kinds. Like the classic version of the descriptive theory, however, it faces the problem of construing the meaning and reference of natural kind terms exclusively in terms of concepts *qua* descriptions in our "heads." Indeed, unless the cluster theory is construed as allowing specialist knowledge in its descriptions, it too will be defeated by the Twin-Earth example (since XYZ satisfies all non-theoretical descriptions).[6]

Most contemporary philosophical theories of the meanings of natural kind terms are mixed causal-descriptive theories. Like (purely) causal theories of reference, these hybrid theories reject the idea that the concepts associated with natural kind terms supply necessary and sufficient conditions for their application. Instead, reference is determined through the interplay (different versions tell different stories) of description and causation. Perhaps the best-known successor to the classic version of the descriptive theory is Chalmers and Jackson (2001)'s two-dimensional semantics. Two-dimensional semantics explicitly rejects the identification of concepts with descriptions, and along with this, the idea that natural kind terms typically have "definitions" that fully determine their extensions (pp. 320–322); the identification of the extension of a term requires "sufficient empirical information about the actual world" (p. 323). Boyd's (1999a, b) homeostatic property cluster theory of natural kinds is a well-known successor to the original version of the cluster theory. Like two-dimensional semantics, it is a hybrid causal-descriptive theory, requiring an "accommodation between conceptual and classificatory practices and causal structures" (p. 406); natural kinds are identified with homeostatic property clusters, more specifically, loose collections of properties, many (but not all) of which tend to co-occur in a significant number of cases as a result of underlying causal mechanisms and processes.

The important point for our purposes is that hybrid causal-descriptive theories follow causal theories of reference in rejecting the idea that the reference of natural kind terms is determined exclusively by concepts; causal relations between language and the world play a crucial role in fixing the membership of the extensions of natural kind terms. As a consequence, these theories provide no more support for a definitional approach to understanding natural kinds than do causal theories of reference. Indeed, the only philosophical theory of the meaning of natural kind terms that is consistent with the definitional approach is the classic version of the descriptive theory, and as we have seen, it was defeated by Putnam's Twin-Earth argument. In this context, it is worth noting that attempts to salvage a "definitional" approach by weakening the concept of definition (e.g., Boyd) are *ad hoc* and misleading,[7] particularly in light of the fact (mentioned earlier) that most philosophers and scientists still couch their definitions of 'life' in terms of necessary and sufficient conditions, or failing that, aspire to do so. Most importantly, emphasizing the construction of definitions ignores what is actually involved in achieving a scientific understanding of a natural kind such as water or life, and even worse, misleadingly suggests that the answer is to be found in an analysis of human concepts, however inadequate they might be.

---

[6] I am grateful to Graeme Forbes for pointing this out to me.

[7] Boyd (e.g., 1999b, pp. 70–72) speaks of homeostatic property clusters "defining" natural kinds but explicitly rejects the idea that they do so by means of necessary and sufficient conditions or, for that matter, on the basis of concepts alone. He is thus using the term 'definition' in a nonstandard way. The important point for our purposes is that the pertinent homeostatic property clusters are not determined solely by means of an analysis of the concept that we associate with a natural kind term, and thus the use of the term 'definition' is misleading.

Putnam's Twin-Earth argument leaves open the possibility that some natural kind terms are such that the concepts associated with them provide necessary and sufficient conditions for their application.[8] This possibility is also left open by other alternatives to the classic descriptive theory. So it seems that one cannot rule out the possibility that 'life' is a rare natural kind term whose extension is fully determined by concepts.

This is possible, but not likely. Our concept of life has changed in major ways since ancient times (Lange 1995). For a long time, self-movement was thought to be a "defining" feature of life. From ancient times until as late as the seventeenth century, a debate simmered over whether stars or planets were alive on the grounds that they are self-moved and their motion resembles that of birds or fish. Moreover, until the mid-nineteenth century it was not clear that fungi should be classified as living since they do not display what were then taken to be essential features of life, namely, movement, growth, and reproduction. Indeed, the "fungus-stone," which resembles a large gray stone and is so tough that a saw is required to cut it, led many early scientists to classify fungi as minerals (Ainsworth 1976, p. 32). The nineteenth-century discovery (by Pasteur and others) that infectious diseases are caused by microscopic organisms (bacteria) that are invisible to the naked eye represents yet another profound change in our concept of life; in the middle ages infectious diseases were attributed to such things as bad air, supernatural influences, and humoral imbalances. The nineteenth century also witnessed the formulation of Darwin's revolutionary theory of evolution, with its emphasis on the historical development of life, through the process of natural selection, which contrasted with the previous emphasis on understanding life a historically in terms of individual organisms. With the advent of molecular biology in the late twentieth century, our concept of life changed yet again, becoming closely coupled to the molecular composition and architecture of organisms—to the presence of a complex cooperative arrangement between carbon-containing macromolecules. It is hardly a coincidence that NASA scientists currently favor the "chemical Darwinian definition" (Joyce 1995). To put the situation into stark perspective, the idea of a fifteenth-century biologist coming up with the chemical Darwinian definition is no more plausible than the idea of a fifteenth-century chemist coming up with a definition of 'water' as $H_2O$.

Unlike the present-day chemical concept of water, however, the biological concept of life will almost certainly continue to change in unanticipated ways. For modern biology lacks the equivalent of a chemical theory of life to delimit the possibilities. In the context of modern chemistry, the unique association of water with the molecular compound $H_2O$ allows scientists to explain and predict properties of water both here on Earth and elsewhere in the universe. These properties include being a good solvent, expanding when frozen, and remaining a liquid over a wide range of temperatures. Many of these properties were thought to be distinctive of water before the advent of molecular theory, and the ability of molecular theory, coupled with considerations from chemical thermodynamics (e.g., Gibb's law), to explain them played a significant role in the acceptance of the claim that water is $H_2O$ (Needham 2002). The association of water with $H_2O$ also allows scientists to explain and predict atypical (otherwise

---

[8] I am grateful to Rob Rupert for pointing this out to me.

unexpected) behavior of water under chemical and physical conditions (very high pressures and temperatures) that do not naturally occur on Earth. This helps to explain why contemporary scientists believe that they have a good understanding of the universal nature of water, and hence a compelling scientific answer to the question 'what is water?'

But as the ongoing disagreement over the "definition" of life reveals, biologists are not at all certain about the possibilities for life. Indeed, as discussed earlier, there are compelling biological reasons for thinking that life as we know it on Earth today is not representative of life in general. The fact that biologists cannot confidently constrain the possibilities for life strongly suggests that our present-day concept of life does not contain identifying knowledge of the kind. In other words, even supposing that some natural kind terms can be successfully defined, there are compelling biological reasons for thinking that 'life' is not among them.

## 3 The quest for a universal theory of life

### 3.1 Scientific theories and theoretical identity statements

The most scientifically compelling answers to "what is" questions about phenomenal natural kinds are provided by scientific theories. The claim that water is $H_2O$ provides a good example. It explicates the nature of water in terms of theoretical entities (hydrogen and oxygen) and processes (covalent chemical bonding) supplied by molecular theory. Other examples of theoretical identity statements are 'sound is a compression wave', 'lightning is an electrical discharge', and 'frogs are amphibians'. In each case, a seemingly natural phenomenal kind (sound, lightning, frog) is explicated in terms of the resources of a widely accepted scientific theory (viz., wave mechanics, electromagnetism, and vertebrate biology).

Despite their grammatical form, most theoretical identity statements do not represent authentic identities. Somewhat ironically, 'water is $H_2O$', the poster child for theoretical identity statements, is one of them. A single molecule of $H_2O$ lacks temperature and pressure, which is a problem since the unique triple point of phenomenal water, a distinctive combination of pressure and temperature at which its gas, liquid, and solid phase coexist in equilibrium, distinguishes it from all other phenomenal chemical substances. Moreover, many of the distinctive phenomenal properties of water (e.g., remaining a liquid over a wide range of temperatures, expanding when frozen, being a good solvent) result from secondary structures (dimers, trimers, and hydrogen bonded networks) produced by "weak" hydrogen bonds among $H_2O$ molecules. Indeed, it is because temperatures and pressures on Earth are close to the triple point of water that we are familiar with many of its distinctive properties; there are many planets where water would not exhibit them. Finally, the "purest" samples of phenomenal water are not composed of just $H_2O$ molecules. For molecules of $H_2O$ invariably dissociate into the ions $H^+$, $OH^-$, and $H_3O^+$ when combined together. The point is while phenomenal water is predominately $H_2O$ it cannot be said to be literally identical with $H_2O$, and most importantly, outside the context of chemical theory, reference to $H_2O$ is not enough to explain and predict the properties of water that distinguish it as a unique kind at the level of phenomenal experience; one needs recourse to the rest of the theoretical apparatus of chemical theory. In short, it is the resources of a scientific theory considered as a whole (with its rich ontology of theoretical entities, properties, mechanisms, and basic "laws") that best explain why a putative natural kind picked out in common discourse possesses the phenomenal properties and displays the overt behavior that it does. It follows that theoretical identity statements cannot be interpreted as "defining" the natural kinds subsumed by scientific theories; among other things, they are not authentic identities, and hence cannot be said to supply (even tentatively) necessary and sufficient conditions for the phenomenal categories concerned. (Nor can they be said to encapsulate scientific theories as a whole because scientific theories invariably systematize and integrate many different phenomenal kinds within their theoretical frameworks and they do not do it in a piecemeal fashion; rather than a theory of water we have a complex theory of chemistry, which subsumes many chemical substances in addition to water.)

Scientific theories suggest previously unrecognized connections among ostensibly unrelated, observable phenomena in virtue of determining which predicates are projectable (may be inductively extended to unexamined members of a class of individuals) and which collections of individuals constitute natural kinds (Psillos 1999; Boyd 1999a, b). It is thus hardly surprising that empirical discoveries made within the context of a new scientific theory sometimes change our ideas about phenomenal kinds in fundamental ways (Cleland and Chyba 2007; Cleland 2006). Indeed, one cannot expect the theoretically refined categories carved out by our best scientific theories to be in perfect alignment with our pre-scientific categories. We might discover that what we thought was a single natural kind consists of more than one natural kind. The term 'jade' provides a good illustration. It was once thought to designate a single natural kind in virtue of the phenomenal properties (color, crystalline form, etc.) exhibited by certain minerals. With the advent of molecular theory, however, it became evident that (despite their similarity in phenomenal properties) these minerals consist of two different molecular compounds. As a consequence, mineralogists split the old category "jade" into two new categories, "jadeite" and "nephrite." Alternatively, we might discover that what we thought were several natural kinds are actually the same natural kind. A good example is the discovery that (despite their striking phenomenal differences) sapphires and rubies have the same molecular composition; as a consequence, mineralogists now classify both as "corundum." It is important to be clear about the nature of such decisions. They are not the product of conceptual analysis, nor do they represent mere verbal decisions (a.k.a. stipulative definitions) to talk about the world in a different way. They represent empirically informed decisions to bring phenomenal categories, picked out in natural language, into closer correspondence with the world as characterized by an empirically well grounded scientific theory; see Boyd (e.g., 1999a,1999b) for a plausible account of how this might work.

The upshot of this discussion is that a scientifically compelling answer to the question 'what is life?' presupposes an empirically adequate general theory (vs. definition) of life. Even supposing that life is, as some (e.g., Keller 2002) have argued, *not* a natural kind, we need a compelling scientific theory to convince us that this is so; for as discussed above, scientific theories play crucial roles in "revealing" the natural kind structure of observable phenomena.[9] In this context, it is also important to keep in mind that we cannot rule out the possibility that life both is a natural kind and lacks a microstructural essence. Microstructural essences are more appealing on purely causal theories of reference than they are on hybrid causal-descriptive theories, which may in part explain Putnam's and Kripke's infatuation with them. But insofar as the classification of observable phenomena into natural kinds is theory dependent (which is widely recognized by both realists and anti-realists), one cannot rule out the possibility of a powerful future theory of life classifying systems as "living" vs. "nonliving" on theoretical grounds other than microstructure. This would hardly be unprecedented. In the general theory of relativity gravity is not a microstructural kind—it is a geometrical property of space-time—which has long posed a problem for unifying general relativity with quantum theory. Years of scientific research directed at identifying "gravitons" have been unsuccessful, and although string theory holds forth the promise of a microstructural account of gravity (and thus a "theory of everything"), it thus far lacks independent empirical support. The point is it is not obvious that a fruitful universal theory of life will provide a microstructural account of the distinction between living and nonliving systems, although of course it may. It is even possible that life is, as Bedau and others claim, an intrinsically functional as opposed to compositional or structural kind. But it is a mistake to conclude (as they seem to do) that this is something that can be decided at the present time with the limited example of life at our disposal. In order to resolve this issue we need additional examples of life.

## 3.2 Searching for alternative forms of life

But how can one search for alternative forms of life given that our experience with life is limited to a single, possibly unrepresentative, example? The problem is exacerbated when one considers that microbial life is probably much more common in the universe than large, complex organisms such as trees and mammals (Ward and Brownlee 2000). The tools currently used to explore the microbial world—microscopy, cultivation, and PCR amplification of rRNA genes—could not detect even modestly different forms of microbial life (Cleland and Copley 2005; Cleland 2007). Put succinctly, in order to formulate a truly general theory of living systems we need unfamiliar forms of life, and yet in the absence of such a theory it is unlikely that we will recognize them as such if we encounter them. In this context, it is important to keep in mind that computer (software or hardware) models of life do not, as sometimes alleged, provide us

---

[9] Although it is beyond the scope of this article, the same could be said for those (e.g., Dupré 1993) who advocate pluralism about taxonomic kinds in biology (which is different from advocating that living systems, as distinct from nonliving systems, do not form a natural kind). Our inability to settle upon a stable taxonomy across different biological subdisciplines, such as organismal biology, population biology, ecology, microbiology, and genetics, might merely reflect the lack of an overarching (unifying) theory of biology, somewhat akin to the state of physics before Newton recognized the commonality in the motion of projectiles, tides, freefalling objects, swinging pendulums, and celestial objects. This is not of course to claim that there are such kinds in biology. It is merely to point out that our inability to identify them does not provide good evidence that they do not exist; the best evidence for pluralism about biological kinds would be a theory of biology that explains why there are no such kinds, just as molecular theory explains why jade is not a natural kind.

with convincing examples of alien life, for their status as living is founded upon the supposition that certain features of familiar biological life are true of all life, and this is just what is at issue.

I have argued elsewhere (Cleland and Chyba 2007; Cleland 2006) that the solution to this conundrum is to search for anomalies. In the context of our concerns in this article, anomalies are physical systems resembling familiar Earth life in provocative ways and yet also differing from it in important and unanticipated ways. As Thomas Kuhn (1962) emphasized, anomalies provide the best grist for the theoretical mill, challenging us to move beyond our preconceptions and consider possibilities that might otherwise be ignored. Kuhn believed that evidence could not be recognized as anomalous within the framework of a well-entrenched theory.[10] But this is not obvious, and even if it were true, it would not apply to life. For as illustrated by the lack of consensus on one of the many "definitions" of life presently being entertained by scientists and philosophers, we currently lack such a theory. Anyone who disagrees is merely presupposing, in the absence of an adequate empirical foundation, that some favored characteristic of familiar biological life (Darwinian evolution, carbon based biochemistry, etc.) is universal to life. Instead of following a definitional approach and searching for life that resembles familiar Earth life in a favored way, we need to search for physical systems that both resemble familiar life in striking ways and also deviate from it in provocative ways. In a nutshell, I am suggesting that scientists jump start the fitful process described by Kuhn whereby anomalies become slowly recognized as such and deliberately search for them.

But how does one search for anomalies? Any investigation must be grounded in our current understanding of biological life on Earth since it provides our only secure example of life. The basic idea behind my proposal is to utilize selected features of familiar Earth life as tentative criteria for life. Insofar as they are construed as tentative, such criteria are understood to be incomplete and, most importantly, defeasible. They are not viewed as either defining (classic version of descriptive theory) or even delimiting (cluster version) life. Indeed, given our current epistemic situation, it is possible that we will someday discover that all the characteristics currently thought to be essential to life are little more than potentially unreliable symptoms of more fundamental but as yet unknown properties. The purpose of tentative criteria for life is not to settle the issue of whether a weird physical system discovered on Mars, for instance, is alive, but rather to focus scientific attention on suspicious physical systems—to identify the best candidates for further scientific investigation.

Unlike alternative definitions of life, tentative criteria do not compete with each other. They can be jointly employed in a search for novel forms of life. It is important that these criteria include a diversity of disparate features of familiar life since we do not know which characteristics of familiar life are essential to all life, wherever it might be found, as opposed to being the result of mere physical and chemical contingencies on the early Earth at the time of the origin of life. Furthermore, tentative criteria need not be universal to familiar Earth life. Features that are common only to life found in certain kinds of environments could prove more useful

---

[10] Actually, Kuhn talked in terms of "paradigms" and later "disciplinary matrixes," both of which encompass more than theories, but this is not important for our purposes.

for searching for life in analogous extraterrestrial environments even if they are not universal to Earth life. Similarly, features of Earth life that are non-existent or very uncommon among nonliving physical systems on Earth could make good criteria for searching for unfamiliar forms of life even if they are not universal, because they stand out against a background of nonliving processes. Thus, for instance, the tiny (4–100 nm), chemically pure, prismatic magnetite crystals found (in the rims of mysterious carbonate globules) in a famous Martian meteorite (ALH84001), recovered in 1984 in Antarctica, were initially cited as providing compelling evidence of fossilized Martian microbes; see Jakosky et al. (2007,pp. 374–378) for a more detailed discussion. This claim was grounded in their striking similarities to magnetite crystals produced by a particular strain of magnetotactic bacteria (MV1) on Earth. MV1 lives in chemically stratified, marine environments. It produces magnetite crystals from dissolved iron, and uses chains of them to orient to its environment. The discovery of strikingly similar crystals in ALH84001 was provocative because at the time no one knew of an abiotic mechanism for producing tiny magnetic crystals of such chemical purity and uniform geometrical shape under natural conditions; indeed, there is still controversy about whether such a mechanism exists, and even if it does, it is not clear that it can explain the crystals in ALH84001. Significantly, their estimated age coincides with a period (around 3.9 billion years ago) in which Mars is thought to have been wet and geologically active, with a powerful magnetic field. If such crystals cannot be (or are rarely) produced abiotically, then, even though they are not universal to life on Earth, they could provide useful criteria for exploring certain extraterrestrial environments for suspicious (anomalous) physical systems.

Some will find my proposal frustrating because it does not answer the question ‘what is life?’ Indeed, it suggests that it might be a long time before we have an answer. There is an important advantage of my proposal over its competitors, however, and that is that it decreases the likelihood of misidentifying an example of truly “alien” life as a nonliving physical system on the basis of a defective definition. This is not a trivial concern. NASA’s use of a metabolic “definition” of life in the 1976 Viking missions to Mars may already have blinded us to an unfamiliar form of life.

The Viking missions are the only dedicated search for extraterrestrial life that has thus far been conducted. Of particular interest to us is the “labeled release” (LR) experiment. One of three metabolic experiments performed robotically on Mars by the Viking lander, the LR experiment yielded results that initially seemed positive for life but quickly turned baffling; for a review of all three experiments, see Klein (1978). Martian soil introduced into the test chamber was injected with a radioactively labeled nutrient solution and started evolving radioactively labeled $^{14}CO_2$—just what one would expect from Earth microbes. When the Martian soil sample was subsequently heated to 160°C for three hours, more than enough to kill Earth microbes, the reaction stopped, strongly suggesting that the initial response had been biological. But when another Martian soil sample was given a second helping of nutrients the anticipated burst of new activity from hungry Martian microbes not only failed to occur, but even more mysteriously $^{14}CO_2$ left over from the initial reaction began disappearing. Scientists were flummoxed. They were facing a genuine anomaly. While closely resembling what one would expect from Earth microbes, the results obtained by the LR experiment also deviated in baffling ways.

Because the Viking experiments were explicitly designed around an officially sanctioned (chemical metabolic) definition, the results of the LR experiment were interpreted as negative; the failure of the Viking gas chromatograph mass spectrometer (GCMS) to detect any organic molecules was treated as conclusive. Non-biological explanations for the mysterious Viking results have since been developed and the current consensus is that they were produced by unusual states of iron. To this day, however, there is no empirical evidence that the Martian surface is strongly oxidizing and, most importantly, NASA currently has no plans for testing this hypothesis in future missions. This illustrates the difference between searching for familiar life and searching for anomalies. Because they did not conform to the favored definition of life, the results of the LR experiments were attributed to a non-biological oxidant, despite the fact that life is also an oxidant. If the Viking experiments had been designed as a search for anomalies, reaction to the LR experiment would have been quite different. It would have been interpreted as just what it was: an ambiguous result worthy of further empirical investigation.

**Acknowledgements** I wish to thank Rob Rupert, Graeme Forbes, and Veronica Vaida for helpful discussions. This work was supported in part by a NASA grant to the University of Colorado's Astrobiology Center.

## References

Ainsworth, G. C. (1976). *Introduction to the history of mycology*. Cambridge: Cambridge University Press.

Bedau, M. A. (1998). Four puzzles about life. *Artificial Life, 4*, 125–140.

Benner, S. A. (1994). Expanding the genetic lexicon: Incorporating non-standard amino acids into proteins by ribosome-based synthesis. *Trends in Biotechnology, 12*, 158–163.

Benner, S. A., Ricardo, A., & Carrigan, M. A. (2004). Is there a common chemical model for life in the universe?. *Current Opinion in Chemical Biology, 8*, 672–689.

Boyd, R. (1999a). Homeostasis, species, and higher taxa. In R. Wilson (Ed.), *Species: New interdisciplinary essays* (pp. 141–185). Cambridge: Cambridge University Press.

Boyd, R. (1999b). Kinds, complexity and multiple realization. *Philosophical Studies, 95*, 67–98.

Chalmers, D. J., & Jackson, F. (2001). Conceptual analysis and reductive explanation. *The Philosophical Review, 110*, 315–360.

Cleland, C. E. (2006). Understanding the nature of life: A matter of definition or theory?. In J. Seckbach (Ed.), *Life as we know it* (pp. 589–600). Dordrecht: Springer.

Cleland, C. E. (2007). Epistemological issues in the study of microbial life: Alternative terran biospheres?. *Studies in History and Philosophy of Biological and Biomedical Sciences, 38*, 847–861.

Cleland, C. E., & Chyba, C. F. (2002). Defining 'life'. *Origins of Life and Evolution of the Biosphere, 32*, 387–393.

Cleland, C. E., & Chyba, C. F. (2007). Does 'life' have a definition?. In W. T. Sullivan & J. A. Baross (Eds.), *Planets and life* (pp. 119–131). Cambridge: Cambridge University Press.

Cleland, C. E., & Copley, S. D. (2005). The possibility of alternative microbial life on earth. *International Journal of Astrobiology, 2*, 165–173.

Davies, P. C. W., & Lineweaver, C. H. (2005). Finding a second sample of life on earth. *Astrobiology, 5*, 154–163.

Dupré, J. (1993). *The disorder of things*. Cambridge: Harvard University Press.

Jakosky, B. M., Westall, F., & Brack, A. (2007). Mars. In W. T. Sullivan & J. A. Baross (Eds.), *Planets and life* (pp. 357–387). Cambridge: Cambridge University Press.

Joyce, G. F. (1995). The RNA world: Life before DNA and protein. In B. Zuckerman & M. Hart (Eds.), *Extraterrestrials—where are they?* (Vol. II, pp. 139–151). Cambridge: Cambridge University Press.

Keller, E. F. (2002). *Making sense of life* (pp. 284–294). Cambridge: Harvard University Press.

Klein, H. P. (1978). The Viking biological experiments on Mars. *Icarus, 34*, 666–674.

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

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lange, M. (1995). Life, 'artificial life,' and scientific explanation. *Philosophy of Science, 63*, 225–244.

Lavoisier, A. L. (1783). On the nature of water and on experiments which appear to prove that this substance is not strictly speaking an element but that it is susceptible of decomposition and recomposition. *Observations sur la Physique, 23*, 452–455 (trans. Carmen Giunta. http://webserver.lemoyne.edu/faculty/giunta/laveau.html).

Locke, J. (1689). An Essay Concerning Human Understanding. Oxford: Oxford University Press.

Mellor, C. H. (1977). Natural kinds. *British Journal for the Philosophy of Science, 28*, 299–312.

Needham, P. (2002). discovery that water is $H_2O$. *International Studies in the Philosophy of Science, 15*, 205–226.

Putnam, H. (1975). The meaning of 'meaning'. In K. Gunderson (Ed.), *Language, mind and knowledge: Minnesota studies in the philosophy of science* (Vol. VII, pp. 131–193). Minneapolis, MN: Minnesota University Press.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. London: Routledge.

Roberts, G. (1994). *The mirror of alchemy*. Toronto: University of Toronto Press.

Ruiz-Mirazo, K., Pereto, A., & Moreno, J. (2002). A universal definition of life: Autonomy and open-ended evolution. *Origins of Life and Evolution of the Biosphere, 34*, 323–346.

Ward, P. D., & Brownlee, D. (2000). *Rare Earth*. New York: Copernicus.

Woese, C. R. (2004). The archaeal concept and the world it lives in: A retrospective. *Photosynthesis Research, 80*, 371–372.

Zemach, E. M. (1976). Putnam's theory on the reference of substance terms. *The Journal of Philosophy, LXXIIII*, 116–126.