From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Philosophy

2017

From Representation to Reality

Hirsch, Robert

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp018g84mp91r

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

FROM REPRESENTATION TO REALITY

Essays on the Nature of Scientific Exegesis

Robert J. Hirsch

A DISSERTATION PRESENTED TO
THE FACULTY OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY
THE DEPARTMENT OF PHILOSOPHY

Advisors: Gideon Rosen and Hans Halvorson

June 2017

© Copyright by Robert J. Hirsch, 2017.
All rights reserved.

For my parents, Rob and Sandy

ABSTRACT

This dissertation examines the nature of scientific exegesis, i.e., the task of determining what science says about the world. I argue that nearly all contemporary theories of scientific exegesis are untenable. My primary argument for this centers around a phenomena that I call 'representational artifacts'. Representational artifacts are, to a first approximation, features of scientific representations that do not directly correspond to elements of the physical world. After exploring the difficulties that such artifacts raise in chapter two, the remaining chapters consider what representational artifacts reveal about the nature of scientific exegesis.

CONTENTS

Abstract iv
Table of Contents v

1 Introduction **1**

2 Scientific Exegesis and Representational Artifacts: A Case Study **6**
1 Maudlin's Mathematical Maneuver 7
2 Metaphysics or Mere Mathematics? 13
3 The Metaphysical Implications of Path-Relativity 29
4 The Challenge of Interpretation 40

3 The Exegetical Import of Representational Artifacts **43**
1 The Problem of Representational Artifacts 43
2 Scientific Literalism 49
3 Ramsificationism 57
4 Exegetical Angst or Interpretive Insight? 66

4 Language, Structure, and Scientific Exegesis **77**
1 Individuation and Abstraction 79
2 Worldly Facts or Linguistic Truths? 86
3 The Tractarian Thesis 94
4 Recasting The 'Linguistic' Objection 96

Bibliography **100**

INTRODUCTION

"But facts are not everything—at least half the business lies in how you interpret them!"

— Fyodor Dostoevsky, *Crime and Punishment*

One test for the depth of a philosophical issue is what, for lack of a better term, we might call the Showing Versus Saying Test: Does the issue exhibit readily recognizable features that are difficult to describe? If so, there's likely something philosophically interesting about the topic. Modality? Normativity? Personal identity? Space and time? All these familiar phenomena are incredibly hard to pin down. And I don't just mean this in the sense that philosophers disagree about *the nature of* any of these topics (they do, of course). Rather, I mean that these topics are hard to pin down in the sense that a good deal of the philosophical work on them revolves around clarifying the very topic enough that we can have any sort of debate. In other words, these topics seemingly put us in a bind: Without knowing what you're talking about, it's hard to debate its nature; yet without knowing its nature, it's hard to know to what you're debating.

Scientific exegesis, I submit, passes the Showing Versus Saying Test. It is relatively easy to explain if explained roughly — that is, when explained with healthy doses of arm flailing, back-of-the-napkin drawing, and their other mnemonic ilk — but difficult to pin down in any precise way. Here's the rough version: Scientific exegesis is the task of determining what the world is like according to science. But this answer is hardly satisfactory. Allow me to illustrate.

Though we have an intuitive feel for the difference between equations like $X = yz$ and $F = ma$, it's difficult to say what the salient difference between them amounts to beyond merely suggesting that the former is a bit of bare mathematical formalism while the latter is a statement of a physical law. But this only relocates the question, for what separates bits of bare formalism from full-fledged physical laws? The obvious suggestion is that the difference amounts to the fact that $F = ma$ *is interpreted* while $X = yz$ is not. But the trouble with this suggestion is that it doesn't say a whole lot. After all, just what does it mean to say that $F = ma$ is interpreted while $X = yz$ is not? Obviously, it doesn't help say that each of the symbols in the former are interpreted. We might instead try to explain the difference in terms of a correspondence between representations and reality: What it means to say that $F = ma$ is interpreted is to say that its constituents (i.e., its symbols) *correspond with* elements of the physical world. But this is problematic for at least two reasons. First, even if the correspondence answer is correct at some level, it cannot be the final answer to the question of what it is to say that $F = ma$ is interpreted. Obviously there are not only many different *ways in which* $F = ma$ could correspond to the world and varieties of physical phenomena to *correspond with*, it's also true that the equation $X = yz$ does, in fact, correspond with elements of the physical world. The trouble here, in other words, is that correspondences are cheap. For the present answer to be satisfying, therefore, it must at a minimum flesh out the details of the correspondence relation. So the first problem is simply that the correspondence answer is, at best, incomplete. But the second and more pressing problem it faces is that it simply gets the exegetical facts wrong in the sense that it just wrongly identifies the content of scientific representations. The foundational difficulty, I will argue in chapters one and two, is that the correspondence answer fails to distinguish between the significant aspects of a representation and those that are mere artifacts. The effect is that, in cases where a representation contain artifacts, the correspondence answer incorrectly judges that these artifacts correspond to genuine aspects of reality and hence leads to misunderstands about the content of such representations.

Let me be clear about what I am suggesting here. I am *not* arguing for some skeptical thesis to the effect that representational artifacts frequently derail our understanding of science. I simply do *not* believe this to be the case. It would be absurd to think that scientists are massively confused about which aspects of their own representations are significant and which are mere artifacts. Instead, I am going to argue that the lesson of The Problem of Representational Artifacts is that scientific exegesis is best understood as factoring into several distinct phases. Moreover, I am going to suggest that The Problem of Representational Artifacts aids in differentiating epistemic and metaphysical variants of the foundational metaexegetical question. The epistemological variant is the question *of determining* where the boundary between the superficial and substantive features of scientific

representations lies. The metaphysical side of the issue, on the other hand, concerns the more foundational question of what it is to be a representational artifact.

I go on to argue that the trouble with extant accounts of scientific exegesis is that they rest on a misconception about the nature of scientific representation that I call The Tractarian Thesis. To a first approximation, it is the thesis that a representation's *structure* constrains what it can represent. My contention will be that, like the correspondence answer, The Tractarian Thesis is viciously circular. The trouble arises from the fact that scientific representations have *infinitely many types* of structure. As a result, we need to know which types of structure are relevant to The Tractarian Thesis. Yet the answer to this depends on the structure of what's being represented. So the basic circularity is (roughly) this: The Tractarian Thesis tells us that R can represent P only if R has the same structure as P; however, which of R's structures is relevant to this question depends on whether R in fact represents P.

In passing, I should say a few words about my choice of terminology. I've chosen the term 'scientific exegesis' rather than 'scientific interpretation' for several reasons. First, although one understanding of the term 'interpretation' would be quite fitting, the trouble is that the term can also mean a number of radically different things. In particular, scientists are sometimes described as interpreting physical phenomena, the outcomes of experiments, and so forth. For example: "By *Faraday's interpretation, the experiment had shown*…that magnetic forces…ran in continuous loops all the way through the magnet."[1] When we speak of the interpretation of an experiment, physical phenomena (e.g., an interpretation of the fossil record), etc., we tend to be putting forward *an explanation or theory*. In such cases, the notion of interpretation involved is a relation that runs upward, from worldly phenomena to the abstract realm of theory.

But the conception of interpretation we'll be concerned with differs significantly. Consider the question Faraday poses to Maxwell:

> When a mathematician engaged in investigating physical actions and results has arrived at his own conclusions, *may they not be expressed in common language as fully, clearly, and definitely as in mathematical formulae?* If so, would it not be a great boon to such as well to express them so—*translating them out of their hieroglyphics* that we might also work upon them by experiment?[2]

This is a paradigm question about scientific exegesis: it concerns the physical significance of Maxwell's equations; that is, it is a question about how to descend from a scientific representation to the physical world. This notion of interpretation is a relation that runs down-

[1] Forbes and Mahon (2014), 67. My emphasis.
[2] Jones (1870), 392. My emphasis.

ward, from the abstract realm of theory to the concrete terra firma of physical phenomena. And it is this sense of interpretation that we'll be concerned with.[3] Given this, the basic first-order interpretive question might be put like this: Given a scientific representation R, what (if any) physical significance does R have? That is, what would the world have to be like to be as R says it is? So I've chosen to use the term 'scientific exegesis' to ensure that we keep this question distinct from questions involving other senses of interpretation.

A second reason for this choice is that 'exegesis' evokes an image of an active task, an image of setting aside surface level details and reading between the lines in search of a text's deeper significance. The term 'interpretation', by contrast, often calls to mind little more than an abstract picture of drawing lines between the scientific lexicon and, in Faraday's words, the common language. But one of my central contentions is going to be that there

[3] While these two notions of interpretation are conceptually distinct, it is important to keep sight of the fact that rarely is there a clean divide between them in practice—a fact illustrated by one Faraday's major discoveries. Throughout the decade following the realization that electric currents produce magnetic forces Faraday came to suspect that the converse was true, that a magnetic force could somehow produce an electric current. Faraday tested the hypothesis by passing a magnet through a coil of wire. An interesting thing happened: the needle on his galvanometer immediately sprang to life, but then, just as quickly, fell still once again. Electricity not only could produce magnetism, but somehow magnetism could also produce electricity. But how? The mere presence of a magnetic force evidently was not the answer, for the current died off once the magnet had been placed within the coil. Faraday's crucial observation came when he pulled the magnet from the coil and the needle again sprang to life, this time leaping in the opposite direction. Intrigued, he proceeded to pass the magnetic back and forth through the coil. Lo and behold, the needle came alive as long as he moved the magnet. Faraday had not only shown that a magnet could produce an electric current, but specifically that *changes in* magnetic force, not just the presence of such a force, were responsible for the flow of current—a fact known today as ***Faraday's Law of Induction***.

While this is a fine summary of the experimental data, the question remains: What does it reveal about the nature of electricity and magnetism themselves? A bit more background is necessary. Over the course of the decade, Faraday had come to adopt a radically new ontology according to which lines of force were 'woven' together into a giant invisible mesh that blanketed the entire universe. One central advantage of postulating this precursor to the electromagnetic field was that it did away with the need for instantaneous action at a distance, long the bane of the Newtonians. For whereas the Newtonians had to maintain that the electric or magnetic force F present at a location ℓ resulted from far away objects acting instantaneously upon ℓ, Faraday could give an explanation in entirely local terms: F was the product of the lines of force passing through ℓ, its magnitude being proportional to the number of lines of force passing through ℓ—the denser the concentration of lines of force at ℓ, the greater F's magnitude at ℓ. It was in light of this ontological picture that Faraday put forward an explanation of induction. They key was the fact that an electric current only arose if the wire or the magnet *were moved*. For as Faraday saw it, when the wire was moved, it 'cut through' the lines of force emanating from the magnet; conversely, when the magnet was moved, the lines of force it produced moved with it 'slicing through' the wire. So on Faraday's interpretation, electric current was generated in a wire when it 'cut through' a magnetic line of force. Thus, Faraday's Law may be formulated as:

Faraday's Law (lines of force formulation) Whenever the number of lines of magnetic force passing through a closed circuit is changed there is an electro-motive force round the circuit represented by the rate of diminution of the number of lines of force which pass through the circuit in the positive direction.

Now, when we speak of 'Faraday's interpretation of electric induction', we are speaking (mostly) about an explanation or theory put forward of some observable phenomena. But even so, what we must observe is that some trace of the present representation-to-world conception of interpretation can be found even here. For in saying that Faraday explained electric induction in terms of lines of force we are presupposing that terms like 'line of force' have already been interpreted—interpreted in the representation-to-world and not the explanatory sense of interpretation. So, in practice, the line between the two notions of interpretation is often blurry. Despite this, from a conceptual point of view, the two notions are distinct.

is far more to articulating the physical significance of scientific representations than such mere lexical correlation. Rather, I'm going to argue that it involves 'reading' science critically, often requiring us to 'read between the lines' of scientific representations to determine their physical significance. So a second reason for my choice of the term 'scientific exegesis' is the hope that it will evoke a more accurate mental image of the task at hand.

2

SCIENTIFIC EXEGESIS AND REPRESENTATIONAL ARTIFACTS: A CASE STUDY

> "*The numerical interpretation…is however necessary…So long as it is not obtained, the solutions may be said to remain incomplete and useless, and the truth which it is proposed to discover is no less hidden in the formulae of the analysis than it was in the physical problem itself.*"
>
> — Joseph Fourier, *The Analytic Theory of Heat*

Recent developments in fundamental physics undermine the foundations of contemporary metaphysics. Or, at least, so it's been argued: According to Tim Maudlin, what are known as 'gauge theories'—a broad family of physical theories such as classical electromagnetism, general relativity, and many quantum field theories including the Standard Model of particles physics—entail a peculiar form of metaphysical relativism where fundamental physical properties are relativized to spacetime paths. For instance, there is no simple fact of the matter about whether Alice's laptop has the same mass as Bob's: their masses may agree relative to one path and yet disagree relative to another. And this variety of metaphysical relativity allegedly precludes the notion of 'similarity' or 'resemblance' central to much of metaphysics. And this leads Maudlin to conclude that gauge theories call for the "wholesale revision"[1] of large swaths of contemporary metaphysics.

While Maudlin's metaphysical picture is provocative and worth discussing—what would path-relative metaphysics really be like?—I would like to focus not on the conclusions that he draws, but instead on his route to these conclusions. For the distinctive feature of his argument is not that it purports to show that fundamental physics poses a threat to meta-

°This chapter was originally published as Hirsch (2015).

[1] Maudlin (2007), 102.

physics. After all, physics is replete with results that have been alleged to have such consequences, e.g., the relativity of simultaneity, quantum non-locality, and so on. Rather, the striking feature of Maudlin's argument is its *overtly mathematical character*: He argues that gauge theories have a revisionary metaphysical character *in virtue of representing physical reality with mathematical structures known as* '*fiber bundles*'. But *prima facie* this inference is puzzling, for a conspicuous gap separates the mathematical character of its premises and the metaphysical nature of its conclusion. This inference will be our principal concern in this paper, for it raises a question of much broader importance: *How is a scientific theory's mathematical formalism related to its ontology?*

The Plan. In section 1, I will begin by presenting an intuitive overview of fiber bundles before turning to consider why Maudlin takes them to call for radical revisions to contemporary metaphysics. Then, in sections 2 and 3, I will turn to evaluating Maudlin's argument. Finally, in section 4, I will draw several lessons about the relationship between the mathematics and metaphysics of scientific theories in general.

1 MAUDLIN'S MATHEMATICAL MANEUVER

So what is the fiber bundle formalism? And why does Maudlin take it to have implications for the ontology of gauge theories and metaphysics more generally?[2]

1.1 AN ILLUSTRATED INTRODUCTION TO FIBER BUNDLES

Fiber bundles are mathematical spaces that one encounters in topology and differential geometry. *Very* roughly, the idea is that they are mathematical spaces 'built' by 'taping on' copies of a space, called a '*fiber*', onto a different space called a '*base space*'. The '*fiber bundle*' itself is the higher-dimensional space comprised of copies of the fiber 'taped on' to each point in the base space.[3]

This is very abstract, so an example will help. Imagine a circular racetrack and a race car driving clockwise around it. Fiber bundles provide an elegant way to represent the race car's journey around the track. To build our fiber bundle, let's start by taking a circle for our 'base space' and copies of the number line as our fibers (see figure 2.1). The base space represents the racetrack while the fibers represent possible speeds of the race car (in the

[2] By 'ontology' I specifically mean '*physical* ontology'. In this paper, I am concerned specifically with the relationship between the mathematics employed by scientific theories and the *non-mathematical* ontology of those theories—e.g., with the relationship between fiber bundles and the physical properties described by gauge theories, like 'charge' and 'flavor'. So I am *not* asking a question in the spirit of the mathematical nominalism debate, e.g., about whether a theory is ontologically committed to mathematical entities in virtue of employing numbers to make calculations. So hereafter I will use 'ontology', 'metaphysical implications', and so on specifically to mean *non-mathematical* ontology, metaphysical implications, etc.

[3] For a precise definition, see Lee (2009), 257 ff.

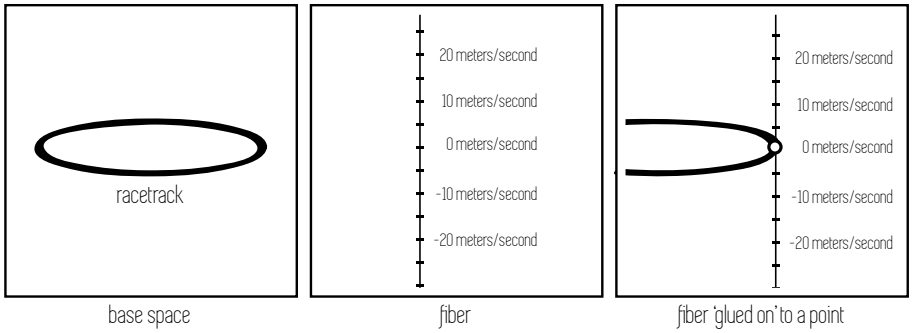


Figure 2.1

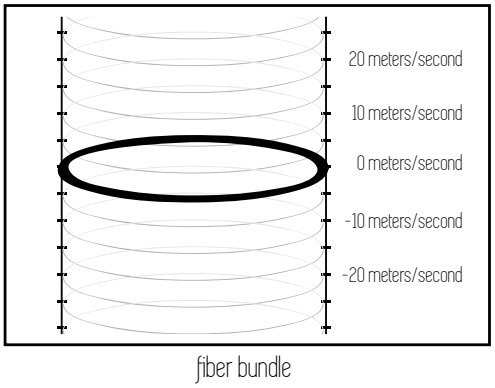


Figure 2.2

clockwise direction). Thus, the fiber bundle itself is a two-dimensional space that 'looks' like a cylinder (see figure 2.2). One nice thing about this formalism is that each point in the fiber bundle encodes multiple pieces of information. In our case, each point in the fiber bundle represents a position on the race track along with a possible speed of the race car at that position. Taking the idea a step further, a curve through the fiber bundle then represents the race car's speed as it travels along a length of the track (see figure 2.3). Our racetrack fiber bundle is (roughly) a particular type of fiber bundle called a 'tangent bundle', however there are many other types of fiber bundles.

Once we've constructed a fiber bundle it's natural to wonder how to compare elements of distinct fibers—after all, we need to be able to make comparisons if we want to determine whether, say, the race car was going the same speed around the first and second turns. Here's where things get interesting. For reasons that we won't go into here, it turns out that one can only compare elements of fibers that are infinitesimally to each other. To compare elements

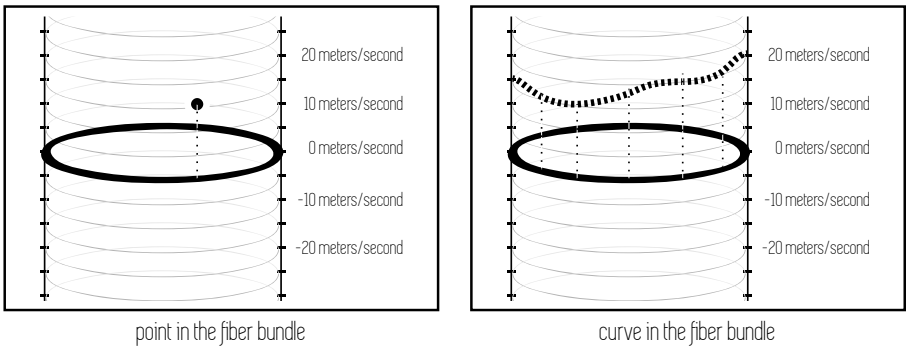


Figure 2.3

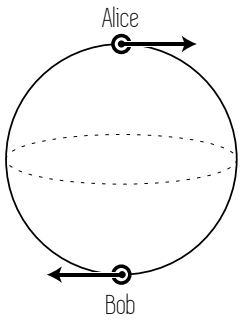


Figure 2.4

of fibers at distant points, one has to 'add up' the infinitesimal comparisons along some paths between these points.

To illustrate, imagine two intrepid explorers, Alice and Bob, standing at the north and south poles, respectively (figure 2.4). From the perspective of the fiber bundle formalism, we *cannot* ask whether Alice and Bob are facing the same direction *simpliciter*. Instead, we can only ask whether Alice is facing the same direction, for instance, before and after taking a very small step towards the south pole.[4] In a derivative way, then, we can ask whether Alice and Bob are facing the same direction by considering whether they are facing the same direction when Alice arrives at the south pole *via some path where her direction stayed the same at each step*. But there's a catch: we will get different answers depending on which path Alice takes to the south pole. For example, if Alice were to walk along the first path then we would say that she was facing the same direction as Bob. If she took the second path, however, then we would say that they were facing different directions (see figure 2.5). In

[4]Strictly speaking, we can't even ask this if we only use the formal machinery discussed so far. Rather, to make local comparisons we need what is called a 'connection'. See section 2 below.

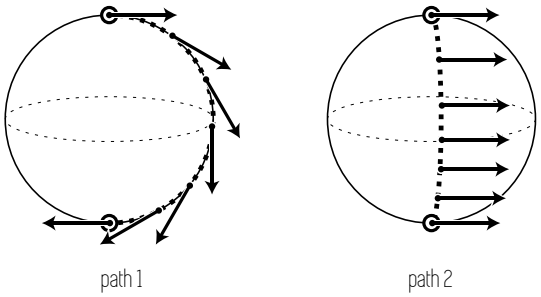


Figure 2.5

other words, comparisons of directions are inherently *path-relative*. (To save ink, I will often shorten locutions like '*comparisons of* property F are path-relative' to 'F is path-relative'.)

1.2 MAUDLIN'S ARGUMENT

While fiber bundles are intrinsically interesting from a mathematical point of view, as Maudlin sees it, their philosophical significance lies in the fact that they are employed by gauge theories where they're used to represent fundamental physical properties. While he intends for his argument to apply to gauge theories at large, Maudlin proceeds by way of considering the example of *quantum chromodynamics (QCD)*.

His discussion of QCD is informal and ultimately boils down to three observations. First, QCD is a gauge theory and, as such, can be formulated using fiber bundles. Second, QCD is part of the Standard Model of particle physics; in particular, it is the part that describes the strong force governing interactions between quarks. And the Standard Model, at least at the time of writing this article, is the best known contender for the theory of fundamental physical reality. Third, to describe these interactions, QCD's physical ontology includes properties known as 'color charges'—properties of quarks that for purposes of the present discussion can be thought of as properties akin to electric charge, though there are three types of color charge, unlike the familiar two varieties of electric charge. From these observations there are three broad steps in Maudlin's argument to the conclusion that large swaths of contemporary metaphysics are in need of wholesale revision.

Step 1. The first step of Maudlin's argument connects mathematical and spatiotemporal notions of path-relativity. To do this, he draws an analogy between color charge and directions on a sphere: "But chromodynamics is a gauge theory, which means that at base *color charge is treated completely analogously to directions*."[5] And from here, it is a short step to the conclusion that color charge is path-relative: "And just as for directions, the results of

[5] Maudlin (2007), 94. My emphasis.

the comparison will in general depend upon the particular path chosen: there is no path-independent fact about whether vectors [representing color charges] in different fibers are 'the same' or 'different'."[6] Thus, from the fact that QCD employs fiber bundles to represent quark color, Maudlin concludes that comparisons of quark color are path-relative. In short, the central claim of the first step is: *In virtue of representing fundamental physical properties with path-relative mathematical features of fiber bundles, gauge theories maintain that fundamental physical properties are spatiotemporally path-relative.*

Step 2. But at least on its own, the mere fact that some physical properties are spatiotemporally path-relative is no cause for a wholesale revision of contemporary metaphysics. After all, no one is tempted to think that a massive overhaul of metaphysics is called for by the mere fact that comparison of directions on a sphere, for instance, are path-relative.

So why, then, is it metaphysically significant that properties like color charge are path-relative? It is widely believed that not all physical properties are created equal. Most, like the property of *either being green or being within 1879 feet of 1879 Hall*, are highly gerrymandered. They mark out no deep, significant feature of reality. They do not, as it is sometimes put, 'carve nature at its joints'. Yet, not all properties are cobbled together. There are an elite few—universals, tropes, sparse properties, and so on—that do, in fact, mark out significant aspects of reality. Their pattern of instantiation explains why, say, Marx Hall resembles 1879 Hall more than it resembles Karl Marx's beard. And theirs is the work of explaining facts about *objective resemblance*. This widespread belief about the metaphysics of properties is often accompanied by a further belief about the aims of science, as David Lewis writes:

> Sharing of [sparse properties] makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are ipso facto not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy.
>
> Physics has its short list of 'fundamental physical properties': the charges and masses of particles, also their so-called 'spins' and 'colours' and 'flavours', and maybe a few more that have yet to be discovered…What physics has undertaken, whether or not ours is a world where the undertaking will succeed, is an inventory of the sparse properties of this-worldly things. Else the project makes no sense.[7]

The view, therefore, is that there are a small number of physical properties that serve to explain objective similarities and dissimilarities, and that the business of science—physics in

[6] Maudlin (2007), 96.
[7] Lewis (1986), 60.

particular—is to uncover which properties enjoy membership in this elite group. In honor of Davids Lewis and Armstrong, two of the most prominent defenders of this view in recent memory, let us call this '*the Lewis–Armstrong picture*'.

It is against the backdrop of this picture that spatiotemporal path-relativity allegedly gets its metaphysical bite. For observe: If F is spatiotemporally path-relative, then two objects may resemble one another with respect to F relative to one path connecting them and yet fail to resemble one another relative to some other path. Maudlin takes this to mean that there is *no fact of the matter* whether the two objects resemble one another with respect to F. In particular, then, if color charge is spatiotemporally path-relative, it cannot explain facts about the objective resemblance between quarks.[8] As Maudlin puts it:

> If we adopt the metaphysics of the fiber bundle…then we must reject the notion that quark color is a universal, or that there are color tropes which can be duplicates, or that quarks are parts of 'natural sets' which include all and only the quarks of the same color, for there is no fact about whether any two quarks are the same color or different.[9]

But if a property fails to explain facts about objective resemblance, then it is not cut out to be a fundamental property according to the Lewis–Armstrong picture. Hence, either color charge is not a fundamental property or the picture must be revised. So, given QCD's present status, Maudlin argues that it is the latter.

In short, the central claim of the second step of the argument is this: *In virtue of representing physical reality with the fiber bundle formalism, gauge theories maintain that fundamental physical properties are spatiotemporally path-relative and hence cannot be accommodated within the Lewis–Armstrong picture.*

Step 3. The last step in the argument draws specific consequences for the Lewis–Armstrong picture and metaphysics more broadly. Maudlin argues that if fundamental properties cannot explain the objective resemblance between objects then the Lewis–Armstrong picture is stripped of its resources for articulating other central notions like that of intrinsic properties, dispositions, and so on. And from here the dominoes begin to fall quickly: For without

[8] It seems to me that the same considerations allow Maudlin to make a much stronger point: Not only are there no facts about the objective resemblance *vis-a-vis* the color charge of distinct quarks, but more strongly, there simply is no fact about the color charge of any quark whatsoever. This is motivated by the fact that the color charge of a quark can fail to resemble itself relative to a path. Suppose that there are two quarks q_1 and q_2 and paths α and β starting at q_1 and ending at q_2 such that the comparison of q_1 and q_2's color charges relative to α differs from the comparison relative to β. In this case, q_1's color charge (similarly with q_2) will differ from itself when compared along the path $\beta^{-1} \circ \alpha$, i.e., the path that results from following α then following β in reverse. But it's a short step from saying that there is no fact of the matter if q_1 resembles itself with respect to color charge to simply saying that there is no fact of the matter about q_1's color charge.

[9] Maudlin (2007), 96.

intrinsic properties, there is no categorical base on which internal and external relations are to supervene.[10] Consequently, there is no duplication. Therefore, no Principle of Recombination. And the list goes on. It is for this reason that Maudlin takes gauge theories to require radical revisions to metaphysics. As he puts it, "since metaphysicians like Armstrong have focused on examples like electric charge and mass in explicating the theory of universals, *eliminating these requires a wholesale revision of that picture of universals.*"[11]

Going forward, our focus will primarily be upon the first two steps of the argument, for these concern the relationship between the mathematics and metaphysics of gauge theories. Step 3, on the other hand, does not directly bear upon this issue and so we will set it aside.

2 METAPHYSICS OR MERE MATHEMATICS?

Our central task in evaluating the first step of Maudlin's argument is to determine if we are licensed to conclude that physical properties like color charge are spatiotemporally path-relative from the fact that they are represented by mathematically path-relative features of fiber bundles. For convenience, call this the '*Path-Relativity Inference*' *(PRI)*.

My evaluation will proceed in two stages. First, in section 2.1, we will consider the status of the PRI within the context of gauge theories at large. We will discover that the PRI is not, in general, valid within the context of gauge theories. Consequently, we will conclude that if gauge theories have metaphysically revisionary implications it is not merely in virtue of the fact that they employ the fiber bundle formalism. This will lead us in section 2.2 to consider how the PRI fares when we restrict our attention specifically to quantum gauge theories, such as QCD. Though we will find that the issues are less clear-cut in this context, we will see that there are a number of reasons that even if we do not outright reject the PRI, we should be wary of arguments that rely upon it.

But first, a few clarificatory remarks. Because Maudlin's argument is informal and does not delve either into the technical details of QCD or the fiber bundle formalism, I will follow suit.[12] As a result, a good deal of the discussion will oversimplify the technical details of gauge theories and the fiber bundle formalism. However, this will do little harm as my objections concern not those technical details, but the philosophical difficulties of extracting

[10] Maudlin uses the terms '*metaphysically pure* internal relation' and '*metaphysically pure* intrinsic property'. For example: "there are no metaphysically pure internal relations. And since internal relations are supposed to be determined by the intrinsic properties of objects, it follows that there are no metaphysically pure intrinsic properties." (Maudlin (2007), 86.) However, these details are unimportant at present as every internal relation/intrinsic property is a metaphysically pure internal relation/intrinsic property.

[11] Maudlin (2007), 102. My emphasis.

[12] In fact, the technical details of Maudlin's discussion are largely uninteresting as many of them are inaccurate—a point which Maudlin is upfront about: "The actual mathematics of chromodynamics is being misrepresented here, since the actual gauge group for QCD is $SU(3)$, not $O(3)$, but these details are not important for the fundamental metaphysical morals." (Maudlin 2007, 95, footnote 7.)

metaphysical conclusions from mathematical considerations. (However, there are technical objections to Maudlin's argument.[13]) The question at hand is what metaphysical conclusions can be drawn about the physical ontology of gauge theories from considerations about the mathematical character of the fiber bundle formalism. It should be unobjectionable to anyone but perhaps the most hardened instrumentalist that conclusions about physical phenomena can be drawn from the mathematical structures they are represented with.[14] So our concern is not *whether* we can draw conclusions about physical phenomena from mathematical considerations, but rather with *the nature of the conclusions* that can be drawn.

2.1 PATH-RELATIVITY IN GAUGE THEORIES AT LARGE

Is the PRI valid in the context of gauge theories at large? Maudlin certainly seems to think it is. In this section I would like to argue otherwise. My claim will be that the PRI is not valid in *classical electromagnetism (EM)* which is *the* archetypal gauge theory upon which other gauge theories—including quantum gauge theories like QCD—are modeled. I begin by considering Maudlin's understanding of the fiber bundle formulation of EM. Then I will present an alternate understanding that lacks the metaphysical consequences of Maudlin's. Finally, I will argue that, in light of considerations about scientific practice and the historical development of gauge theories, my understanding is the more plausible of the two.

Now to be clear, I want to stress that our concern here is with the question of what metaphysical/non-mathematical conclusions *can be drawn from* considerations about a theory's mathematical formalism. What we must keep sight of is that the first two steps of Maudlin's argument are driven almost entirely by *considerations about the mathematical character of fiber bundles.* We therefore are not interested at present in whether the conclusion of his argument is true, but only if his argument is valid. And in arguing that it is invalid I am only making the modest claim that there are *at least some* gauge theories where mathematical path-relativity does not imply spatiotemporal path-relativity. Whether these are isolated cases, or whether the PRI is invalid in gauge theories at large, is a question that will remain to be answered. (But even if these are isolated cases, and hence the PRI is valid in most gauge theories, this would at best only allow the first step of the argument to go through. And as I will argue in section 3, even if the PRI were valid there are still serious problems elsewhere in the argument.)

[13] For example, one objection is that Maudlin brushes over the technical fact that gauge theories are not formulated using just any fiber bundles, but rather what are known as '*principal* fiber bundles'—and these add to the abstract nature of the fiber bundle formulation of gauge theories. A more important technical objection is the fact that there are other ways to formulate gauge theories that do *not* employ fiber bundles, e.g., what is known as the 'holonomy formulation'. See Healey (2007), 111–128.

[14] Thanks to Jim Weatherall for pushing me to be clear about this point.

The Fiber Bundle Formulation of Electromagnetism

So how does one express classical electromagnetism in the fiber bundle formalism?[15] Here is how one author puts it: "Classical gauge field theory turns out to be a theory of connections on a principal fiber bundle…"[16] And more specifically: "Electromagnetic field theory is a connection in a circle bundle.…[E]lectromagnetism is a $U(1)$-gauge theory, that is, electromagnetism is the theory of connections on a principal $U(1)$-bundle over the space-time manifold."[17] And another author writes, "Maxwell's theory of electromagnetism is described by the $U(1)$ gauge group."[18] But what does any of this mean? What's a '$U(1)$-bundle' or '$U(1)$ gauge group'? And what's a 'connection'?

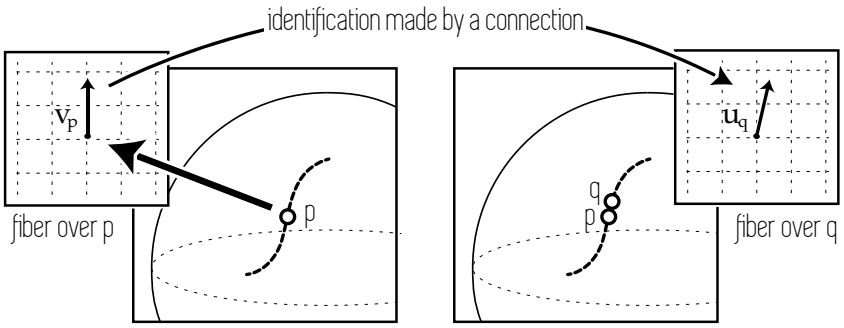


Figure 2.6

Let's start with the notion of a 'connection'. Intuitively speaking, we can think of a connection as a mathematical gadget for comparing elements of distinct fibers in a bundle. Think back yet again to Alice and Bob. A connection can be thought of as a device that tells us what counts as 'the same direction' each time Alice takes a *very* small step towards Bob. A connection therefore allows us to compare the directions Alice and Bob are facing when they are at opposite poles by providing a way to determine whether Alice's direction has changed as she walked to the south pole. In this sense, a connection provides a *local notion of sameness*. More generally (and very roughly), we can think of a connection as a machine that allows us to make comparisons along a path. If we hand the connection an element of the fiber over some point on a path, the connection tells us what to 'identify'

[15] To do justice to the fiber bundle formulation of EM would require far more technical machinery than I have space to develop here—concepts like principal bundles, Lie groups, and so on. Accordingly, I have chosen to err on the side of making the presentation of the material accessible. For an extremely detailed presentation of the fiber bundle formulation of classical electromagnetism see Naber (2011), chapter 2. Other helpful presentations include Akyildiz (1983), section 4.1; Darling (1994), 194–231; Baez and Muniain (1994), 223–241; Nakahara (2003), chapter 10; and Garrity (2015), chapters 18–20.

[16] Akyildiz (1983), 179.

[17] Akyildiz (1983), 219-220.

[18] Nakahara (2003), 399.

that element with in infinitesimally nearby fibers (see figure 2.6). To compare elements of fibers over points that are not infinitesimally nearby, we simply add up all of the infinitesimal comparisons along a curve between the points.

Now what is a '$U(1)$-bundle'? e can roughly think of a $U(1)$-bundle as a fiber bundle whose fibers are copies of a circle. The elements of the individual fibers are points on the circle. And because one wants to represent the electromagnetic field on spacetime when formulating EM, one specifically employs a $U(1)$-bundle whose base space is a four-dimensional manifold. Thus, we can roughly think of a $U(1)$-bundle for EM as a fiber bundle where we've taped a copy of a circle onto each spacetime point.

But what does any of this tell us about EM? That is, how should we physically interpret what connections and $U(1)$-bundles represent? At least two questions must be addressed when interpreting any theory's mathematical formalism. First, which features of the mathematical formalism are physically significant? Second, what is the physical significance of these features? In the present context: Which features of the fiber bundle formalism have significance for the ontology of EM, and what is the ontological significance of those features?

Maudlin's Interpretation

The following passage is representative of Maudlin's interpretation of the fiber bundle formulation of EM:

> Fiber bundles provide new mathematical structures for representing physical states, and hence a new way to understand physical ontology. For example, modern electromagnetic theory holds that what we call the 'electromagnetic field' just *is* the connection on a fiber bundle. Such an account evidently carries with it quite a lot of ontological commitments: there must be a base space, and internal degrees of freedom at each base point represented by a fiber, and a unified object that corresponds to all of the fibers with a connection.[19]

Maudlin's answer to the first interpretive question is clear: according to his understanding of the fiber bundle formulation of EM, the physically significant mathematical features include the $U(1)$-bundle, its base space, each of its fibers, and the connection on the bundle. What is less clear, however, is Maudlin's answer to the second question, but it seems to be this: there would have to be a physical entity represented by the base space of the fiber bundle, a different entity represented by the fiber bundle itself, and a still further entity that corresponds to the connection. But this leaves something to be desired. One wants

[19] Maudlin (2007), 101. Emphasis in the original.

to know not only which mathematical features of the fiber bundle formalism are physically significant, but also what their physical significance is. What are these entities? What, for instance, is this 'physical connection'?

Maudlin does not say. For as Maudlin sees it, such questions are misguided as they attempt to shoehorn the ontology of modern physics into traditional metaphysical categories. Instead, on his view, the line between mathematics and physical reality is, at least methodologically speaking, intentionally thin. Continuing the previous passage:

> But if one asks whether, in this picture, the electromagnetic field is a substance or an instance of a universal or a trope, or some combination of these, none of the options seems very useful. If the electromagnetic field is a connection on a fiber bundle, *then one understands what it is by studying fiber bundles directly, not by trying to translate modern mathematics into archaic philosophical terminology.*[20]

Yet despite eschewing explanations of the physical significance of fiber bundle formalism that are couched in non-mathematical terms, Maudlin still takes this formulation of EM to be incompatible with traditional metaphysical categories:

> If an electromagnetic field is a connection on a fiber bundle, then there are more things in heaven and earth than dreamt of in Plato's or Aristotle's philosophy. And surely this is to be expected: it would be a miracle if the fundamental ontological structure of the universe fit neatly in the categorical pigeonholes handed down to us from two millennia ago.[21]

Indeed, it is in this context that Maudlin says that the fiber bundle formalism calls for a "wholesale revision"[22] of the Lewis–Armstrong picture.

An Alternative Interpretation

There is, however, an interpretation of the fiber bundle formulation of EM that does *not* have these metaphysical implications.[23] Briefly, as I see it—as do many others—the fiber bundle formulation of EM is a notational variant of a more well-known formulation of EM that is understood to have a familiar physical ontology consisting of a four-dimensional spacetime, a unified electromagnetic field, and physical quantities like mass and electric charge—physical quantities that are *not* path-relative.

[20] Maudlin (2007), 101. My emphasis.
[21] Maudlin (2007), 101.
[22] Maudlin (2007), 102.
[23] My interpretation draws heavily on the resources mentioned in footnote 15.

In vector calculus, one represents the electric and magnetic fields with vector fields,

$$\mathbf{E} = (E_x, E_y, E_z) \qquad \text{and} \qquad \mathbf{B} = (B_x, B_y, B_z),$$

that assign vectors encoding the electric and magnetic field strengths to each point in the manifold representing physical space—think of **E** and **B** as attaching arrows to locations that indicate the direction and magnitude of the electric and magnetic forces that a particle would experience if it were placed at those locations. But treating the electric and magnetic fields as separate physical objects has the untoward consequence that the strength of the magnetic field varies from one reference frame to another.[24] To resolve this, we need to think of the electric and magnetic fields as a single physical object: the electromagnetic field.[25] But mathematically, how do we represent such an object? Enter what we might call the 'tensor formulation' of EM. From a mathematical perspective, the unified electromagnetic field can be represented by a *tensor*, $F_{\mu\nu}$.[26] Think of a tensor as a machine that 'eats' vectors and spits out real numbers, where the number of Greek indices indicate how many vectors the tensor consumes—e.g., $F_{\mu\nu}$ eats two vectors. The important thing to notice is how $F_{\mu\nu}$ encodes **E** and **B**:

$$F_{\mu\nu} \;=\; \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}.$$

Essentially all we've done is rewrite the components of **E** and **B** using a matrix. But unlike these vector fields, $F_{\mu\nu}$ provides a mathematical way to represent the fact that the character of the electromagnetic field does not depend upon one's reference frame.[27]

Now there is a mathematical trick that allows us to express the equations describing a vector or tensor field in a more convenient format—a trick that will turn out to be crucial for understanding the fiber bundle formulations of EM and gauge theories more generally.[28]

[24] This is one of the main motivations for special relativity.

[25] We also need to adopt a Minkowski spacetime.

[26] More precisely: a tensor field.

[27] Exactly what it means for something not to depend upon one's reference frame depends upon the particular theory one is considering. In the present context, this should be understood to mean that the electromagnetic field behaves covariantly under the Lorentz transformations.

[28] Locally speaking—that is, in sufficiently small regions of a manifold—this trick is always available to us (this is the content of what is known was the 'Poincaré lemma'). However, if a space is not what is known as 'simply-connected' then there will be vector/tensor fields with no corresponding global potential function. The paradigm case is the tensor (1-form)

$$\theta \;=\; \frac{x\,dy - y\,dx}{x^2 + y^2},$$

on the 'punctured plane', i.e., $\mathbb{R}^2 \setminus \{0\}$. There is no scalar function ω (0-form) such that $d\omega = \theta$. The question of which vector/tensor fields have corresponding potentials is studied by a branch of mathematics known as 'de Rahm cohomology'.

Let's warm up by considering a more concrete example. To compare the altitudes of some locations, the easiest way is to proceed indirectly by introducing a conventional point of reference, e.g., sea level. Doing so not only simplifies calculations, but it also reduces the amount of information we have to keep track of—e.g., rather than independently keeping track of the differences in altitude between Everest and K2, K2 and Denali, and Denali and Everest, we need only keep track of each mountain's altitude relative to sea level. Vector and tensor fields can be represented in an analogous fashion using what are called 'potential functions' (or simply 'potentials'). In philosophical jargon, potentials can be thought of as convenient mathematical ways for encoding *relational* facts. For instance, when EM is formulated using vector calculus, **E** can be encoded by an 'electric potential' while **B** can be expressed with a 'magnetic potential', typically denoted by V and $\mathbf{A} = (A_x, A_y, A_z)$ respectively.[29] And working with the potentials V and **A** rather than with the vector fields **E** and **B** affords conveniences analogous to those in the case of altitude—e.g., there are fewer functions to keep track of, computations are simpler to perform, and so on. Potentials play a similar role in the tensor formulation of EM. However, there is an important difference: Because we are representing a unified electromagnetic field, we no longer want to have separate electric and magnetic potentials; rather, we want to use a single 'electromagnetic potential', typically denoted by A.[30] Note that just as $F_{\mu\nu}$ encodes **E** and **B** in a single object, A encodes the electric and magnetic potentials into a single potential function:

$$A = -V dx^0 + A_x dx^1 + A_y dx^2 + A_z dx^3.$$

Despite the convenience of potentials, there is a catch: the numerical values of a potential are *not* physically significant, only *differences in values of the potential* are. One physics text puts it like this:

> Evidently potential as such carries *no real physical significance*, for at any given point we can adjust its value at will…In this sense it is rather like altitude: If I ask you how high Denver is, you will probably tell me its height above sea level, because that is a *convenient and traditional reference point*. But we could as well agree to measure altitude above Washington D.C., or Greenwich, or wherever. That would add (or, rather, subtract) a fixed amount from all our sea-level readings, but it *wouldn't change anything about the real world. The only quantity of intrinsic interest is the difference in altitude between two points, and that is the same whatever your reference level.*[31]

[29] The bold-faced **A** indicates that the magnetic potential is a vector-valued function. By contrast, the electric potential is a scalar function.

[30] Unfortunately, the standard notation is somewhat confusing. To be clear: A_x, A_y, and A_z are the components of the magnetic potential, V is the electric potential, and A without subscripts is the unified potential function for the electromagnetic field that is often called the 'electromagnetic four potential'.

[31] Griffiths (1999), 80. My emphasis.

This is crucial to keep in mind as we consider how to understand the tensor and, ultimately, fiber bundle formulations of EM.

With this in mind, let us turn to considering the metaphysical picture painted by the tensor formulation of EM. Let's begin with the first of our two interpretive questions: Which mathematical features of the tensor formulation are physically significant?[32] The standard answer is that the physically significant features of the tensor formulation include the four-dimensional manifold, the tensor $F_{\mu\nu}$, as well as the results of performing various operations on the potential A.[33] The potential itself, however, does *not* correspond to a real physical property. As for the second question—what is the physical significance of these features?—standardly, these mathematical features are taken to represent a four-dimensional spacetime, a unified electromagnetic field, and physical properties like mass and electric charge. Moreover, the mathematical features of the tensor formulation constrain how these elements of physical reality relate to one another. The relationship between force, mass, and charge on the one hand, and the unified electromagnetic field on the other, is constrained by $F_{\mu\nu}$ roughly as follows: $F_{\mu\nu}$ represents the force that the electromagnetic field would exert on a particle p given p's mass, velocity, position, and electric charge.[34]

This metaphysical picture differs dramatically from the one painted by Maudlin's interpretation of the fiber bundle formulation of EM. In the first place, Maudlin claims that the fiber bundle formulation eliminates physical properties like electric charge and mass: "since metaphysicians like Armstrong have focused on examples like *electric charge and mass* in explicating the theory of universals, *eliminating these* requires a wholesale revision of that picture of universals."[35] By contrast, the tensor formulation of EM does *not* eliminate properties like electric charge and mass. In fact, as we have just seen, the physical significance of the tensor $F_{\mu\nu}$ is described in terms of properties like electric charge and mass. Moreover, on the tensor formulation, some fundamental physical properties, such as electric charge, simply are not spatiotemporally path-relative—it makes sense to ask, *simpliciter*, if the electric charge of particles at different locations are the same.

Now clearly the tensor and fiber bundle formulations of EM differ in many ways. For example, whereas the latter assigns tensor values to points in a four-dimensional manifold, the former employs a connection on a higher-dimensional manifold. And given these formal differences, one might expect that there would also be significant metaphysical differences between what the two formulations represent. However, as I will now argue, the two

[32] Warnick and Russer (2014) provides a detailed discussion of the physical significance of the tensor formulation of EM.

[33] For example, taking the exterior derivative of A gives us the tensor that represents the electromagnetic field: $dA = F$.

[34] For a more detailed discussion see Misner et al. (1973), 71–74.

[35] Maudlin (2007), 102. My emphasis.

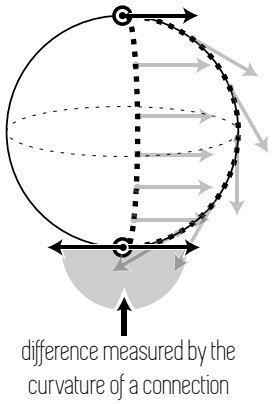


Figure 2.7

formulations are just different ways of expressing the very same physical content. The formal differences do *not* indicate substantive differences in physical ontology but rather are merely matters of notational variation. Thus, if I am correct, the metaphysical implications of the fiber bundle formulation are no different than those of the tensor formulation.

Above, I said that in the fiber bundle formulation, the electromagnetic field is represented with a connection on a $U(1)$-bundle. This statement is extremely imprecise. More carefully, it is not quite that the connection represents the electromagnetic field as much that *the curvature of* a connection represents the electromagnetic field strength. Roughly speaking, the curvature of a connection measures how much the comparisons provided by a connection will disagree along different paths. In the case of Alice and Bob, for example, whereas the connection provides a means of comparing Alice's direction along different spatial paths, the curvature of that connection measures the extent to which the comparisons would disagree along different paths between the same points (see figure 2.7).

But it's still quite imprecise to say that the curvature of a connection on a $U(1)$-bundle represents the electromagnetic field strength. What we need to consider is *how* the curvature of such a connection *is used* by physicists to represent the electromagnetic field strength. If one digs into the details, one will find that a connection is just a new way of encoding the potential for the electromagnetic field, and that the connection's curvature is used to encode differences in the potential. In a discussion of classical electromagnetism, Edward Witten puts it like this: "From a modern point of view, the vector potential is a connection, which we will call A... F is interpreted as the curvature of the connection..."[36] And another author writes, "We can now recast Maxwell's equations into the language of con-

[36] Witten (2007), 363.

nections and curvature. The end result will be that *the potential… will be a connection… and the electromagnetic two-form will be the curvature of this connection.*"[37]

But what does this tell us? Recall that potentials are not directly physically significant. Rather, their physical significance lies in *differences in* their values. Similarly, we can think of the choice of a connection roughly as a way of numerically representing some quantities. But as with potentials, the particular 'labels' that a connection assigns are not themselves physically significant. Rather, the important thing is how those labels vary from place to place, for this is what represents differences in electromagnetic field strength:[38]

> Thus we can indeed describe Maxwell's equations as follows: Given a charge density and a current, we have a fixed current one-form J. Among all possible connections…we choose those connections A whose corresponding curvature $F = dA$ satisfies [Maxwell's equations]…The key…is that the field of the force, here the electric and magnetic fields, is described as the curvature of a connection, which we write for emphasis as
>
> $$\boxed{\text{Force} = \text{Curvature}}$$ [39]

In short, the strength of the electromagnetic field is represented in the fiber bundle formulation of EM by the curvature of a connection on a $U(1)$-bundle. In other words, the curvature of a connection on a $U(1)$-bundle is just a fancy way of rewriting $F_{\mu\nu}$!

So according to the orthodox opinion, the fiber bundle formulation of EM is representationally equivalent to the tensor formulation. They both describe a unified electromagnetic field and its interactions with physical entities due to their electric charge. Moreover, connections play the same role in the fiber bundle formulation that potentials play in the tensor formulation.[40] And if this is correct, then the connection does *not* correspond to a physical entity as Maudlin suggests. Rather, the physically significant feature is the *curvature of* the connection. Moreover, given that the physical properties described by the tensor formulation are *not* spatiotemporally path-relative, it follows that the physical properties—e.g., mass, charge, and so on—represented by the fiber bundle formulation are *not* spatiotemporally path-relative.

Now if this is correct, then the fiber bundle formulation of EM lacks the path-relative metaphysical implications suggested by Maudlin. But on the assumption that he is correct in maintaining that the fiber bundle formulation of EM employs path-relative mathematical

[37] Garrity (2015), 264. Emphasis mine.

[38] For a helpful discussion of how the mathematical formalism is used to represent the physical phenomena, see Aitchison and Hey (2003b), chapter 13.3. Also see Garrity (2015), chapters 18–21.

[39] Garrity (2015), 265.

[40] The role of potentials is more complicated in the quantum context, see section 2.2.

features to represent physical quantities, we can therefore conclude that *the PRI is invalid in the fiber bundle formulation of EM, and hence not valid for gauge theories at large.*

Historical Considerations

It remains to be shown, then, that my interpretation of the fiber bundle formulation of EM is more plausible than Maudlin's.

In the first place, the fiber bundle formalism is only one of many mathematical frameworks which EM has been formulated with throughout its history. It has also been formulated using vector calculus, tensor calculus, and differential forms, to name a few.[41] What do actual physicists make of this? Do they take each of these different mathematical frameworks to paint different pictures of EM's physical content? Let's see. Above we saw one author describe the fiber bundle formalism as merely providing a new language for EM: "We can now *recast Maxwell's equations into the language of connections and curvature*."[42] Baez and Muniain (1994) echo this sentiment in a chapter titled "Rewriting Maxwell's Equations" where they use phrases like 'the *language* of differential forms', and at one point even say, "These are just the first pair of Maxwell's equations *in slightly newfangled notation!*"[43] And yet another author writes:

> There thus arises the questions whether it is worth while to learn the language and use the methods of fiber bundles since so far it has been possible to do without them. It is hard to give a straightforward and convincing answer to this question; probably the only reasonable thing to say is "the future will tell." My personal opinion is that at least some concepts of fiber bundle theory will become an established part of mathematical physics because fiber bundles provide a natural and *convenient framework* for discussing the concepts of relativity and invariance, describing gravitation and other gauge fields…*Fiber bundles provide a convenient language* for dealing with local problems of differential geometry and field theory.[44]

Wu and Yang (1975) even present a 'dictionary' for translating the physical concepts of gauge theory into the geometric language of the fiber bundle formalism, and vice versa. Other examples could be cited, but let's not belabor the point. What these examples suggest is that physicists and mathematicians do not view the fiber bundle formalism as painting a radically different picture of EM's physical content. Rather, they view it as merely providing a different language for expressing EM. And in this sense, the fiber bundle formalism

[41] For an overview of a number of other formulations, see Thirring et al. (2003), section 7.3.
[42] Garrity (2015), 264–265. My emphasis.
[43] Baez and Muniain (1994), 71. My emphasis.
[44] Trautman (1980), 287. My emphasis.

is on a par with vector calculus, the tensor formalism, etc. The various formulations simply represent the same physical facts in mathematically different ways. And this strongly suggests that the differences between the tensor and fiber bundle formulations of EM are not indicative of ontological differences but rather are merely notational variations.

Maudlin's interpretation does not sit well with this outlook. Consider the stance he takes toward the fiber bundle formulation in the following passage:

> *Fiber bundles provide new mathematical structures for representing physical states,* and hence a new way to understand physical ontology. For example, modern electromagnetic theory holds that what we call the 'electromagnetic field' just *is* the connection on a fiber bundle. *Such an account evidently carries with it quite a lot of ontological commitments:* there must be a base space, and internal degrees of freedom at each base point represented by a fiber, and a unified object that corresponds to all of the fibers with a connection.[45]

Maudlin seems to be suggesting that the fiber bundle formulation of EM has metaphysical implications that other formulations do not. But if this were the case, then the fiber bundle formulation would *not* be a mere notational variant of, e.g., the tensor formulation. Maudlin may be happy to accept this, but it comes at the price of failing to respect the attitude found in actual scientific practice towards these formulations. By contrast, my understanding of the fiber bundle formulation is in complete agreement with, and in fact is motivated by, the prevailing attitude in actual scientific practice.

Further, it wasn't realized until relatively late into the development of EM and gauge theories that fiber bundles could be used in formulating such theories. For example, when C. N. Yang—one of the two fathers of modern gauge theories—was asked whether it was true that mathematical entities like fiber bundles and connections were an intended part of gauge theories, he answered:

> Totally false. What Mills and I were doing in 1954 was generalizing Maxwell's theory. We knew of no geometrical meaning of Maxwell's theory, and we were not looking in that direction. To a physicist, gauge potential is a concept rooted in our description of the electromagnetic field. Connection is a geometrical concept which I only learned around 1970.[46]

It would seem strange, then, if fiber bundles were an essential part of the content of EM, or gauge theories more generally. A more natural view of the matter is that fiber bundles provide a mathematical framework that proves to be useful for representing the physical

[45] Maudlin (2007), 101. Italics are my emphasis; underlines are emphasis in the original.
[46] Zhang (1993), 15.

content of EM—physical content that had been, and still often is, studied and expressed in many other mathematical frameworks.

A story will help motivate this outlook. Suppose that you've taken great pains to catalog the objects in your office. You've triple- and quadruple-checked that your list is correct. Then I come along and say, "I've examined your list, given it some thought, and decided it would be better expressed in Latin." I hand you my list. That was strange, you think. But then I add, "By the way, according to my list, your office has some interesting objects that you didn't write down. You know how it is—Latin just brings such an interesting ontology with it." You'd wonder, who is this crazy person and what is he talking about? At a minimum, you would think that if my list really was different, it just wasn't an inventory of your office. Maudlin's interpretation of fiber bundles is to EM as my ontologically loaded Latin is to the contents of your office. If the fiber bundle formulation of EM had distinctive ontological commitments that the vector calculus or tensor formulations did not, then one would expect that the reception of the fiber bundle formalism would have been like the reception of my Latin list—it just gets the facts wrong. But historically this simply isn't what happened.

The best explanation for this, I submit, is that fiber bundles and their ilk do *not* introduce any distinctive ontology as Maudlin suggests. Rather, they simply offer new ways to represent a familiar ontology that had been, and continues to be, represented with other mathematical formalisms. Maudlin assumes that by employing new mathematical structures one incurs new ontological commitments and hence infers that the fiber bundle formulation of EM has a distinctive ontology. But as I see it, the inference runs in the opposite direction. The reason fiber bundles can be used to formulate EM is *precisely because fiber bundles don't have any distinctive implications for the physical ontology of EM, rather they are just a new way of formulating the same theory.*

In light of these considerations, I maintain that my interpretation of the fiber bundle formulation of EM is more plausible than Maudlin's. And if this is correct, then the fiber bundle formulation of EM does *not* have the metaphysical implications that Maudlin suggests. And more importantly, it means that *the PRI is not valid within the context of gauge theories at large*.

2.2 PATH-RELATIVITY IN THE QUANTUM CONTEXT

We've seen that representing some part of physical reality with path-relative features of the fiber bundle formalism does not, in general, give rise to the metaphysical implications that Maudlin suggests. Perhaps, however, there is something peculiar about *quantum* gauge theories (QGTs) that lends itself to Maudlin's argument.

On the face of it, QGTs appear friendlier to the PRI than their classical counterparts. Recall that in EM, while potentials are used to represent electromagnetic field strength, the potentials themselves are merely mathematical conveniences that lack any direct physical significance. In the quantum context, however, potentials play a more significant role. For one thing, potentials are used to represent the so-called '*phase*' of a quantum particle. For another, a phenomenon known as the '*Aharonov–Bohm effect*' provides evidence that potentials are not merely mathematical conveniences but rather represent real physical properties.[47] In fact—and this is the central observation that lends support to the PRI in the quantum context—it turns out that the potentials used to represent the phase of quantum particles are not only *mathematically path-relative*, but moreover, the AB-effect also provides evidence that *the phase of a quantum particle is spatiotemporally path-relative*! Thus, not only do potentials don a newfound physical significance in the quantum context, but moreover, it appears that path-relative mathematical features represent spatiotemporally path-relative properties. So it would seem that the quantum context vindicates the PRI!

But let's not be hasty, for these are delicate matters. What we must keep in mind is that our concern lies with *the validity* of Maudlin's argument—that is, with whether we can infer that gauge theories describe spatiotemporally path-relative properties from the fact that they employ path-relative mathematics. I want to suggest that despite the fact that the AB-effect provides evidence that QGTs describe spatiotemporally path-relative properties, *the AB-effect casts doubt upon the validity of the PRI*. The key question that we must consider is this: What determines whether potentials are physically significant? The AB-effect suggests that the answer, whatever it is, must necessarily involve non-mathematical considerations. For at least mathematically speaking, there is no difference between potentials in the classical and some quantum contexts.[48] In philosophical jargon, there is no 'intrinsic' difference between potentials in the two contexts. And yet, potentials are physically significant in one context but merely mathematical conveniences in the other.

[47] *Very* roughly, the AB-effect occurs when a quantum particle passes through a region where the magnetic field strength is zero everywhere but the magnetic flux from a nearby region—a region that the particle does *not* pass through—influences the particle's phase, thereby producing interference patterns. In the standard presentation of the AB-effect, charged particles are shot towards a solenoid—essentially, a long cylindrical coil of wire through which an electric current is passed—but are split into two beams of particles that travel around the outsides of the solenoid and then are recombined on the opposite side. While the magnetic field strength is zero everywhere outside of the solenoid, the particles undergo different phase shifts depending upon which path they travel along, and this produces the observed interference patterns. For helpful introductions to the AB-effect, see Baez and Muniain (1994, 130–140) and Griffiths (2005, chapter 10.2.3); for a more in-depth discussion, especially of the philosophical issues raised by the AB-effect, see Belot (1998) and Healey (2007, chapter 2).

[48] To be clear, I am not saying that, in general, there is no mathematical difference between potentials in classical and quantum contexts. My point here is just that the electromagnetic potential is taken to be physically insignificant in the classical context and yet, due to the AB-effect, physically significant in the context of quantum mechanics. The burden is therefore on Maudlin to show what, if any, mathematical differences there are between potentials in the classical and quantum contexts, and to show why these differences are metaphysically significant.

This suggests that there is a fundamental flaw in the PRI even in the quantum context. What the AB-effect illustrates is that considerations about how a formalism is used are crucial for determining whether some mathematical feature is physically significant. The trouble with the PRI, however, is that it pays little heed to how the fiber bundle formalism is actually used by physicists and in doing so neglects an essential aspect involved in interpreting a theory's formalism. So regardless of whether the conclusion of the PRI happens to be true in the case of QGTs, there is strong reason to suspect that the PRI is invalid given that it fails to consider how the fiber bundle formalism is actually used by physicists.

While this is far from conclusive, a number of other considerations bolster the case against the PRI in the context of QGTs. In the first place, a hallmark of gauge theories is that they they have features known as '*gauge freedom*': roughly, mathematical features that lack any physical significance.[49] One text on the subject even uses this as a working definition: "[Gauge theories] are theories in which the physical system being dealt with is described by more variables than there are physically independent degrees of freedom…"[50] Another puts it like this: "Gauge theories are also deeply disturbing and unsatisfying in some sense: They are built on a redundancy of description."[51] Whenever gauge freedom is a factor, as Maudlin himself has stressed in the context of quantum gravity, one must exercise a great deal of care when drawing conclusions about a theory's ontology from its mathematics, lest—to borrow Maudlin's phrase—we 'abuse gauge freedom to create metaphysical monstrosities'.[52] At minimum, this suggests that the PRI paints an overly simplistic picture of what is involved in interpreting a theory's formalism. For in light of gauge freedom, it is far from clearcut *which* mathematical features of the fiber bundle formalism are physically significant and which are merely formal artifacts. But the PRI pays no heed to this. Instead, it attempts to draw metaphysical conclusions blithely from mathematical considerations without giving any thought to whether the mathematical features in question are physically insignificant.

A second consideration is simply that QGTs are, after all, *quantum* theories. As such, they independently bring along a host of thorny problems, like the infamous 'measurement problem', which make interpreting quantum theories incredibly difficult—as Maudlin himself has been keen to emphasize.[53] In fact, many of the well-known interpretive problems that arise in ordinary quantum mechanics prove to be *more* difficult when we turn to QGTs.[54]

[49] As it happens, potentials in EM are one of the paradigm examples of gauge freedom. And although the AB-effect provides evidence that potentials have a physical significance in the quantum context that they lack in EM, it does *not* provide reason to conclude that there is no gauge freedom in QGTs.

[50] Henneaux and Teitelboim (1992), xxiii.

[51] Zee (2003), 171.

[52] Maudlin (2002a).

[53] For instance, see Maudlin (2002b).

[54] And quantum field theories, more generally. See Barrett (2002).

And the quantum nature of QGTs provides further reason to be wary of the PRI. For even if one identified *which* features of the fiber bundle formalism were physically significant, in light of the quantum nature of QGTs it would still be highly nontrivial to say *what* physical significance these features have.

But the most significant consideration stems from the mathematical state of QGTs. Were QGTs rigorously formulated, then there might be a case to be made that one could investigate the physical content of QGTs solely through an investigation of the fiber bundle formalism as Maudlin suggests, e.g.: "If the electromagnetic field is a connection on a fiber bundle, then one understands what it is *by studying fiber bundles directly*…"[55] However, at the time of writing, no one knows if there are any realistic mathematical models of QGTs. I do not merely mean to say that no one knows if any of the models of QGTs accurately describe the actual world. Rather, I mean that no one even knows what a realistic mathematical model of a quantum gauge theory would even be, let alone if there are any. Arthur Jaffe and Edward Witten put it like this: "[O]ne does not yet have a mathematically complete example of a quantum gauge theory in four-dimensional space-time, nor even a precise definition of quantum gauge theory in four dimensions."[56] Indeed, the issue is so difficult that there is a million dollar prize for showing that a model exists![57] Yet, despite the fact that there is as of yet no rigorous mathematical formulation of QGTs, physicists have not only been successful in using such theories to make extremely accurate predications, but they also take the physical content of QGTs like the Standard Model to be fairly well-understood.

This suggests that 'studying fiber bundles directly' is simply the wrong thing to do if one wants to understand the physical, and ultimately metaphysical, significance of QGTs.[58] If there is some reason for believing that some of the physical properties described by QGTs are spatiotemporally path-relative, the reason is not primarily mathematical in nature. But precisely what the PRI purports to do is provide mathematical reasons for believing that gauge theories describe spatiotemporally path-relative properties. Accordingly, we have further cause to view the PRI as suspect, even in the quantum context.

Unfortunately, QGTs are decidedly more difficult than their classical counterparts and so it is hard to say anything definitive about the PRI. However, I want to suggest that these

[55] Maudlin (2007), 101. My emphasis.

[56] Jaffe and Witten (2006), 132.

[57] This is one of the seven Millennium Problems called the 'Yang–Mills Existence and Mass Gap'. For a precise statement of the problem, see `http://www.claymath.org/millenium-problems/yang-mills-and-mass-gap`.

[58] Indeed, contrast Maudlin's outlook towards the relationship between mathematics and physics with the attitude expressed by Yang in the interview I mentioned above: "**Zhang [the interviewer]:** Is it important for a physicist to learn a lot of mathematics? **Yang:** No. If a physicist learns too much mathematics, he or she is likely to be seduced by the value judgment of mathematics, and may lose his or her physical intuition.…**Zhang:** For a physicist, experimental results are more important to learn? **Yang:** This is right." (Zhang 1993, 20).

considerations cast sufficient doubt upon the PRI such that, if nothing else, we should refrain from employing it when inquiring into the metaphysical implications of QGTs.

To recap, the validity of the first step of Maudlin's argument depends upon an inference that I've been calling the 'Path-Relativity Inference'. I argued that considerations about how path-relative mathematical features *are used* in the context of classical electromagnetism—the archetypal gauge theory—in conjunction with historical considerations, show that the PRI is *not valid* within the context of gauge theories at large. And while I stopped short of claiming that the PRI was also invalid in the narrower context of quantum gauge theories, I argued that there were strong reasons to avoid the PRI when investigating the metaphysical implications of QGTs. But given that the PRI is at the heart of the first step of Maudlin's argument, we must conclude that the first step of the argument does not succeed.

3 THE METAPHYSICAL IMPLICATIONS OF PATH-RELATIVITY

The argument in the previous section turned, at least in part, on the inchoate mathematical state of QGTs. It is important, therefore, to consider what merits Maudlin's argument would have if future developments in QGTs were to support the PRI. What are the metaphysical implications if we assume that the PRI is valid and, consequently, that properties like color charge are spatiotemporally path-relative?

As Maudlin sees it, if physical properties like color charge are spatiotemporally path-relative, then gauge theories are incompatible with the Lewis–Armstrong picture. For recall, the primary work of fundamental physical properties on this picture is to explain objective resemblances between distinct entities. But according to Maudlin, this is work that spatiotemporally path-relative properties are unfit to do. Here is a representative passage:

> If we adopt the metaphysics of the *fiber bundle to represent chromodynamics, then we must reject the notion that quark color is a universal*, or that there are color tropes which can be duplicates, or that quarks are parts of 'natural sets' which include all and only the quarks of the same color, for there is no fact about whether any two quarks are the same color or different.[59]

Later, he qualifies the claim: "The fan of philosophical ontology (universals, tropes, natural sets, or whatever) can no doubt find a way to jerry-rig an object that can do the same mathematical work as a fiber bundle…"[60] For short, let us say that a piece of ontology (e.g., a property, relation, etc.) is '*exotic*' if the Lewis–Armstrong picture either cannot accommodate it, or can accommodate it only through some highly gerrymandered account. In

[59] Maudlin (2007), 96. My emphasis.
[60] Maudlin (2007), 101.

these terms, Maudlin's claim is that gauge theories say that properties like color charge are exotic in virtue of the fact that they represent such properties with the fiber bundle formalism.

But I would like to suggest otherwise. I will argue that even if we assume that the PRI is valid and hence mathematical path-relativity implies spatiotemporal path-relativity, it does not necessarily follow that properties like color charge are exotic.

3.1 EXOTIC ONTOLOGY OR MUNDANE METAPHYSICS?

The question we must consider is: Why are spatiotemporally path-relative properties unfit to explain facts about objective resemblance?

Consider Maudlin's paradigm example: directions on a sphere. Above, we saw that we could represent the directions that our explorers, Alice and Bob, are facing using the tangent bundle to the sphere. We also saw, however, that the elements of fibers in the tangent bundle are path-relative, so assuming that the PRI is valid, it follows that directions are spatiotemporally path-relative. So it is certainly true, for instance, that Alice would be facing a different direction were she to walk along path 1 rather than path 2. On Maudlin's understanding, the physical significance of such path-relativity is that it implies that *there is no fact about* whether Alice and Bob are facing in the same direction when they are standing at distant locations. And if there is no fact about whether they are facing the same direction, then directions cannot be the sorts of properties that feature in comparisons of objective resemblance. Of course, while directions are not serious candidates for fundamental physical properties, Maudlin's claim is that because "color charge is treated completely analogously to directions",[61] the metaphysical significance of path-relativity in the case of color charge is the same as in the case of direction: viz., "…*there is no fact about* whether any two quarks are the same color or different."[62] More generally, let us call this variety of spatiotemporal path-relativity which implies that there 'is no fact about' the similarity between such properties, '*exotic path-relativity*'.

The trouble with this line of reasoning is that it concludes more than is licensed by the PRI. The issue is that exotic path-relativity is only one of many ways to understand the physical significance of spatiotemporal path-relativity. And although the PRI underwrites the inference from mathematical path-relativity to spatiotemporal path-relativity, it provides no guidance about which understanding of the physical significance of spatiotemporal path-relativity is the correct one. For instance, in the case of Alice and Bob, the PRI in conjunction with the mathematical considerations about the tangent bundle imply that Alice and Bob's directions are spatiotemporally path-relative. If understood as exotic path-

[61] Maudlin (2007), 94.
[62] Maudlin (2007), 96. My emphasis.

relativity, the physical significance of this would be that there is no fact of the matter about whether they are facing the same direction.

But this is not the only way to understand path-relativity. There is also what we might call '*mundane modal path-relativity*'. Say that a property F is path-relative in the mundane modal sense if comparisons of F between some objects x and y depend upon the histories of x and y—i.e., x and y agree with respect to F but *would have* disagreed if x had some other history. This captures the natural sense in which directions are path-relative. For instance, *were* Alice to walk along path 1 (see figure 2.5 above) she *would* be facing the same direction as Bob, but *were* she to walk along path 2 she *would not*. This is a far cry from claiming that there is 'no fact of the matter' about whether Alice and Bob are facing the same direction. Indeed, many aspects of ordinary experience are path-relative in this mundane modal sense—e.g., Bob's weight depends upon whether his path goes through the gym or through the pub, whether the fragile vase arrives intact depends on whether there are any sudden drops along its path, and so on. Crucially, observe that mathematical path-relativity in conjunction with the PRI only implies that Alice and Bob's directions are spatiotemporally path-relative in *some* sense, however it does *not* imply what sense this is.

Could the physical significance of path-relativity in gauge theories be understood in the mundane modal sense? A number of things would need to be fleshed out to provide a satisfactory answer, but the basic idea is straightforward enough: to say that color charge is spatiotemporally path-relative in the mundane modal sense is simply to say, for instance, that the colors of some quarks q_1 and q_2 *would* agree after q_1 is moved to q_2's position along one path, but *would have* disagreed if q_1 *had been* moved along some other path. What is important to notice is that, so understood, spatiotemporal path-relativity does not pose a challenge to the Lewis–Armstrong picture. For given the mundane modal understanding, it does not follow from the fact that color charge is spatiotemporally path-relative that there is no fact of the matter about the similarity between two quarks' colors. So, if this were a tenable understanding of spatiotemporal path-relativity, then Lewis and Armstrong could grant Maudlin's premises and yet still deny his conclusion.[63]

[63] One might worry that there is an important difference between directions on a sphere and properties like color charge: unlike directions, color charge is a *fundamental* property. And surely, the worry goes, understanding fundamental properties in such modal terms is anathema to the Lewis–Armstrong picture. However, while it is certainly true that some modal construals of path-relativity would be anathema, it is not so clear that all are. There seems to be an important difference between understanding properties as inherently modal, and understanding the path-relativity of such properties in modal terms. Surely Lewis and Armstrong have no qualms about saying, for example, that if some positively and negatively charged particles *were* placed in close proximity, then they *would* attract each other—from the perspective of the Lewis–Armstrong picture, to say that positive and negative charges attract isn't to say that such properties are themselves modal properties. Rather, it is a second-order claim about the modal profiles of such properties. It seems to me that there would be nothing unorthodox from the perspective of the Lewis–Armstrong picture about understanding spatiotemporal path-relativity as a second-order modal property.

However, whether mundane modal path-relativity ultimately provides a way for Lewis and Armstrong to avoid Maudlin's conclusion is not the important point. Given that our concern is with the validity of Maudlin's argument, the crucial thing is this: Even in conjunction with the PRI, it does not follow from the fact that gauge theories employ the fiber bundle formalism that properties like color charge are exotic. So even if it does not ultimately prove tenable to understand spatiotemporal path-relativity in the mundane modal sense, our discussion has revealed an inferential gap between mathematical and spatiotemporal path-relativity that is not filled by the PRI. Consequently, Maudlin is making a substantive assumption about how the fiber bundle formalism is used to represent physical reality when he claims that it shows that properties like color charge are exotically path-relative.

At a minimum, then, Maudlin needs to provide reasons for believing that exotic path-relativity is the best way to understand the physical significance of mathematical path-relativity. I have no argument that this cannot be done, but allow me to briefly highlight two obstacles. First, because mathematical formalisms do not come, so to speak, with labels attached announcing which of their features are representationally significant—let alone which features are used to represent fundamental properties—there is a considerable amount of interpretive latitude available in understanding a formalism's physical significance. For instance, if the mundane modal understanding proves to be untenable, Lewis and Armstrong could simply deny that path-relative mathematical features represent fundamental properties. Doing so would open the door to understanding spatiotemporally path-relative properties as extrinsic properties, disguised relations, or any number of other things. After all, nothing intrinsic to the fiber bundle formalism implies that path-relative features represent fundamental properties.[64] Second, even if Lewis and Armstrong were to grant that such features represent fundamental properties and so ultimately concede that their picture must be modified, this does not mean that a wholesale revision is in order. After all, the Lewis–Armstrong picture can accommodate a variety of seemingly problematic ontology (e.g., vector quantities) by making minor modifications that do not abandon the spirit

[64] Again, it is important to keep in mind that in evaluating Maudlin's argument, we are considering what follows from considerations about the mathematical character of the fiber bundle formalism. If we broaden our considerations to include how the fiber bundle formalism *is put to use by actual physicists*, then there is far less interpretive latitude for Lewis and Armstrong to exploit in accommodating spatiotemporally path-relative properties.

of the view.[65] Of course, how satisfying these responses are will depend upon details which I do not have space to go into here.

But the point here is a modest one: it is merely that there are a number of strategies open to Lewis and Armstrong for accommodating path-relative fundamental properties that, while requiring some minor modifications, are far from a wholesale revision.

3.2 IT'S ALL RELATIVE?

So it's one thing to claim that the Lewis–Armstrong picture has difficulty accommodating certain pieces of ontology. It is quite another thing to argue that it is in need of wholesale revision. However, it seems to me that Maudlin argues for a stronger claim that to some extent circumvents this consideration. His suggestion is not merely that gauge theories posit *some* exotic properties. Rather, his claim is that they maintain that *all* fundamental physical properties are path-relative and hence exotic. For if this were true, it would seem that no minor modification to the Lewis–Armstrong picture would allow it to accommodate gauge theories. Look at it like this. If all fundamental physical properties are spatiotemporally path-relative, then there are no properties suited for explaining objective resemblance. Consequently, there are no intrinsic properties, no notion of duplication, no Principle of Recombination, and so on. And were this true, it would certainly seem to call for something close to a wholesale revision of the Lewis–Armstrong picture.

A number of passages suggest that Maudlin means to argue for this stronger claim. For instance, consider when he suggests that gauge theories preclude sparse properties: "And if one believes that fundamental physics is the place to look for the truth about universals (or tropes or natural sets), then one may find that physics is telling us *there are no such things*."[66] The natural way to read this passage is *not* as the claim that fundamental physics precludes understanding quark color, though only quark color, in terms of universals. The natural reading, rather, is that physics precludes understanding *any* fundamental physical property in terms of universals. That is, the natural way to understand the claim here is that there

[65] By way of analogy, consider how Humean supervenience might accommodate vector quantities. According to Humean supervenience, all matters of fact supervene on the qualitative features of point-sized regions of spacetime and spatiotemporal relations between them. But vector quantities, like force, require extended regions of spacetime to define. Do vector quantities call for the wholesale revision of Humean supervenience? Almost certainly not. There a number of things that the Humean can say in response that, even if they go against the letter of Humean supervenience strictly understood, are not in tension with the spirit of the view. First, she might simply argue that vector quantities can be analyzed in terms of qualitative properties and spacetime relations. In this case, accommodating vector quantities would not require any modification of her view. But suppose that she comes to believe that such a strategy will not succeed—what then? There are still a number of strategies she may pursue before abandoning her view wholesale. For instance, she may maintain that everything supervenes on *extremely small, though extended, regions of spacetime* rather than on *point-sized regions* of spacetime; cf., Arntzenius (2012). Or she may expand her view of fundamental ontology and take the supervenience base to include not only qualitative features of spacetime points, but also primitive vector quantities; cf., Bricker (1993). These are minor modifications that do not abandon the spirit of Humean supervenience.

[66] Maudlin (2007), 96. My emphasis.

simply are no universals. And if this claim, then Maudlin's argument cannot merely be that quark color is path-relative, but rather must be that *all* fundamental properties are. Or to give a second example, consider when he suggests later that the fiber bundle formalism is incompatible with internal relations: "Duplication is supposed to be a metaphysically pure internal relation between items. But from the point of view of fiber bundle theory, it makes no sense to 'copy' the state of one region of space-time…There is no metaphysical copying relation such as the Principle of Recombination presupposes."[67] Like before, the claim here is not merely that there is no fact of the matter about whether distinct regions are duplicates *with respect to quark color*. The claim, rather, is that there is no fact of the matter about whether the regions are duplicates in *any* respect; there are no facts about the similarity between *any* properties in different spacetime regions. But given the broad contours of Maudlin's argument, the support for such a conclusion comes only from arguing that *all* fundamental properties are path-relative.

If this reading is correct, then Maudlin's claim is this: in virtue of employing the fiber bundle formalism, gauge theories maintain that *all* fundamental physical properties are path-relative. For convenience, I will refer to this claim by saying that gauge theories are '*exhaustively path-relative*'. The question therefore is: *Are gauge theories exhaustively path-relative?*

Now if one digs deeper, one will discover that, in point of fact, gauge theories—even quantum gauge theories—simply do not maintain that *all* fundamental physical properties are spatiotemporally path-relative. For instance, in both EM and QCD, neither mass nor electric charge are spatiotemporally path-relative. And in quantum electrodynamics (a QGT), not only is electric charge path-independent, but it is actually presupposed in putting the fiber bundle formalism to use.[68] So Maudlin's claim that gauge theories are exhaustively path-relative simply is not borne out by the actual physical content of such theories. However, while this indicates that something has gone wrong in Maudlin's argument, it does not shed light on what the problem is. So for the moment, let's set these considerations about the actual physical content of gauge theories aside. For what we want to consider is what reasons there might be for thinking that the fiber bundle formalism shows that gauge theories are exhaustively path-relative.

Unfortunately, while Maudlin is clear about why he takes the fiber bundle formalism to show that color charge is path-relative, he is less explicit about why he takes it to show that *all* fundamental properties are. The closest he comes is in a passage that I partially quoted above:

[67] Maudlin (2007), 103.

[68] Roughly speaking, to represent the physical phenomena described by quantum electrodynamics, one specifies the connection on a fiber bundle with reference to the charge of an electron. See Aitchison and Hey (2003a), chapter 7.4; also Baez and Muniain (1994), 130–140, provides a good overview of the role electric charge plays in interpreting the formalism in the context of quantum mechanics.

> *If we adopt the metaphysics of the fiber bundle to represent chromodynamics*, then we must reject the notion that quark color is a universal, or that there are color tropes which can be duplicates, or that quarks are parts of 'natural sets' which include all and only the quarks of the same color, for there is no fact about whether any two quarks are the same color or different. Further, we must reject the notion that there is any metaphysically pure relation of comparison between quarks at different points, since the only comparisons available are necessarily dependent on the existence of a continuous path in space-time connecting the points. So it seems that there are no color properties and no metaphysically pure internal relations between quarks. *And if one believes that fundamental physics is the place to look for the truth about universals (or tropes or natural sets), then one may find that physics is telling us there are no such things.*[69]

The operative idea here seems to be that if we grant what we might call the '*Completeness Thesis*'—the thesis that physics provides a complete inventory of the world's fundamental physical properties, relations, etc.—then it follows that worlds described by gauge theories are devoid of path-independent fundamental properties.

But this misunderstands the fundamental issue. Once again, we must keep sight of the overtly mathematical character of Maudlin's argument. The argument is that gauge theories are exhaustively path-relative in virtue of employing the fiber bundle formalism. But observe that while this argument draws conclusions about the physical content of gauge theories from mathematical considerations, the Completeness Thesis does not concern the relationship between a theory's mathematics and its physical content. Rather, it concerns *the status of* its physical content. It is a thesis concerning whether the world contains anything beyond the entities posited by physics. The claim that gauge theories are exhaustively path-relative, on the other hand, *is an interpretive claim*. It concerns what gauge theories say that the world is like. The Completeness Thesis is simply orthogonal to this claim. So the question is not, as Maudlin seems to suggest, whether we should infer that all fundamental properties are path-relative from the fact that gauge theories maintain that they all have this character. Rather, the key question—what we might call the 'Exhaustivity Question'—concerns *the interpretation of* gauge theories:

The Exhaustivity Question: What reason is there to think that gauge theories are exhaustively path-relative in virtue of employing the fiber bundle formalism?

[69] Maudlin (2007), 96. My emphasis.

Unfortunately, Maudlin's answer to the Exhaustivity Question is hazy. However, if we are willing to be somewhat speculative, it seems to me that we can discern the threads of two different answers running through his discussion.

The first answer that Maudlin hints at is this: The fiber bundle formalism is simply *incompatible with* the existence of spatiotemporally path-independent properties and therefore we can infer that gauge theories are exhaustively path-relative. This answer is suggested by passages such as: "If we adopt the metaphysics of the fiber bundle to represent chromodynamics, then we *must reject* the notion that quark color is a universal…"[70] However, we can quickly dispatch this answer. In the first place, it is encouraged in no small part by the failure of Maudlin's account to clearly demarcate the physical ontology of gauge theories from the fiber bundle formalism. This failure fosters the impression that the connection between the mathematical formalism and physical content of gauge theories is closer than it actually is. And this, in turn, encourages the idea that mathematical features of the fiber bundle formalism can be straightforwardly incompatible with the existence of various physical properties. For if mathematical features and physical ontology are closely identified, then it is a short step from mathematical incompatibility to conclusions about what an ontology cannot contain. For instance, from claims like "the 'electromagnetic field' just *is* the connection on a fiber bundle",[71] to conclusions such as, the "curvature [of a connection] is not a Humean property…"[72]

But the second, and more serious reason, why this answer is unsatisfactory is that it begs the question. Let us grant that the fiber bundle formalism precludes certain *formal* properties and relations. This simply is not what is at issue, for this concerns the *mathematical* character of fiber bundles. Rather, the question at hand is why we should think that gauge theories are exhaustively path-relative. And this is a question about the *physical* content of gauge theories. Saying that the fiber bundle formalism *precludes some mathematical feature* is one thing, saying that *gauge theories preclude some physical feature* is another thing entirely. Indeed, the Exhaustivity Question is equivalent to asking why we should think that the fiber bundle formalism precludes the existence of spatiotemporally path-independent fundamental properties. So the proposal that fiber bundles are simply incompatible with the existence of path-independent fundamental properties cannot be the answer to the Exhaustivity Question.

The second answer to the Exhaustivity Question that seems to emerge in Maudlin's discussion has the ring of an indispensability argument. He writes: "[M]ost importantly, the physics *does not need* any such absolute comparison of quark colors: the physics *can be stated*

[70] Maudlin (2007), 96. My emphasis.
[71] Maudlin (2007), 101. Emphasis in the original.
[72] Maudlin (2007), 100.

entirely in terms of the connection on the fiber bundle."[73] More generally, the idea appears to be that if some physical phenomena can be represented by a theory without employing some mathematical feature, then that feature is a physically insignificant aspect of the theory's formalism. But how does this answer the Exhaustivity Question? The idea seems to be this: If the Completeness Thesis is true, then it follows that the only fundamental properties are those described by gauge theories. In particular, this means that if gauge theories only describe path-relative properties, then we can infer that all fundamental properties are path-relative. And given that we can conclude that a physical property is spatiotemporally path-relative from the fact that it is represented by a path-relative mathematical feature via the PRI, we therefore have the conditional: If gauge theories represent physical properties exclusively with path-relative mathematical features, then gauge theories are exhaustively path-relative. Hence, the Exhaustivity Question reduces to what we might call the 'Exclusivity Question':

The Exclusivity Question: What reason is there to think that fundamental physical properties are *represented exclusively by* path-relative features of the fiber bundle formalism?

Enter the indispensability step: If gauge theories only need to use mathematically path-relative features, we can infer that no path-independent mathematical features is representationally significant. That is, we can infer that gauge theories represent fundamental physical properties exclusively with path-relative mathematical features. Hence, it follows that gauge theories are exhaustively path-relative.

There are a number of problems with this answer, but allow me to skip ahead to the most serious. The question we must ask is: How does one determine which mathematical features are required for representing some physical phenomena and which features are not? In particular, why should we think that path-relative features are the only mathematical features needed for representing fundamental physical properties? Maudlin never says. More significantly, the present answer, like the first, begs the question. For look at it like this. Either gauge theories *do* or *do not* say that all fundamental properties are spatiotemporally path-relative. Now if we suppose that the reasoning leading to the Exclusivity Question is correct, then there are two cases. On the one hand, if gauge theories *do*, in fact, say that all fundamental properties are spatiotemporally path-relative, then such properties *must be represented exclusively by* mathematically path-relative features of the fiber bundle formalism. On the other hand, if gauge theories *do not* say that all fundamental properties are spatiotemporally path-relative, then such properties *cannot be represented exclusively by* mathematically path-relative features. But which one is it? The trouble here is that to determine whether fundamental physical properties are represented *exclusively by* path-relative math-

[73] Maudlin (2007), 96. My emphasis.

ematical features, we antecedently need to know what gauge theories say about the world. But this was precisely the point of asking whether fundamental properties are represented exclusively by path-relative features in the first place. In short, the second answer to the Exhaustivity Question presupposes the very thing it purports to account for.

While neither of the answers discernible from Maudlin's discussion succeeds in addressing the Exhaustivity Question, it is natural to wonder if perhaps there is *some* compelling reason to think that gauge theories are exhaustively path-relative in virtue of employing the fiber bundle formalism. But there is a deeper *a priori* reason that considerations about the mathematical character of the fiber bundle formalism—even in conjunction with the PRI and the Completeness Thesis—*could not* show that gauge theories are exhaustively path-relative. Namely, that there is no straightforward connection between the *presence or absence of features in a theory's mathematical formalism* on the one hand, and the *existence or non-existence of entities in its physical ontology* on the other. For if the presence of mathematical features were indicative of the existence of physical entities, then we should infer that there are physical entities corresponding to potentials in EM. But as we saw above, while potentials afford a convenient method for describing the electromagnetic field, potentials do not themselves have any direct physical significance. (Though as we saw, in some contexts they do.) Conversely, if the lack of explicit mention of some mathematical features were indicative of the non-existence of physical entities, then we would be led to infer, for instance, that there was no electric field according to the fiber bundle formulation of EM. For in the fiber bundle formulation, a model of EM is presented by specifying a connection on a $U(1)$-bundle despite the fact that the connection itself is not the physically significant mathematical feature. Rather, as we saw, the physically significant feature is the *curvature of* the connection.

The lack of a straightforward connection between a theory's mathematical formalism and its ontology stems from the fact that—to borrow a saying from Cantor—"the essence of mathematics lies in its freedom."[74] That is, there are no set-in-stone rules for how a mathematical formalism must be used to represent some physical phenomena. In particular, the mathematical features *explicitly* mentioned when representing some phenomena need not have any direct physical significance. While conversely, many *implicit* mathematical features may have considerable physical significance. Think again of potentials in EM. As we have seen, while one can present a model of EM by specifying a potential, potentials do not have any direct physical significance. What is physically significant, however, are *differences in* potentials. So while potentials may be *explicitly* mentioned and differences in potentials are left *implicit*, the former are not physically significant while the latter are. It would be interesting to know how far this phenomenon goes, but we do not have time to pursue it here. What it points to, however, is that one has a great deal of flexibility when mathematically

[74] As quoted in Martin (1996), 94.

representing some physical phenomena. And given this, even if we assume the PRI is valid and hence that mathematical path-relativity implies spatiotemporal path-relativity, we only have reason to conclude that path-relative mathematical features *are* physically significant. However, the PRI does *not* provide reason to think that path-independent mathematical features are physically *in*significant. Thus, even assuming that the PRI is valid, we still do not have reason to conclude that gauge theories are exhaustively path-relative.

How, then, can one determine which mathematical features comprise *all and only* the physically significant features of the fiber bundle formalism? In contrast with Maudlin's proposal that "philosophers would do best to try to understand the structure [of fiber bundles] *on their own terms*",[75] I want to suggest that the answer must be this: *One must begin with some antecedent understanding of the physical ontology of gauge theories*, and then consider how the fiber bundle formalism is actually put to use in representing it. For without an antecedent grasp on what a mathematical formalism is used to represent, as we have been discovering, one cannot determine which aspects of a formalism are physically significant and which are merely formal artifacts. And this suggests, more generally, that understanding a theory's formalism 'on its own terms' is necessarily of limited value in understanding the theory's physical content.

So what then of the Exhaustivity Question? Do path-relative mathematical features comprise all and only the physically significant features of the fiber bundle formalism? The answer is no. For as we saw above, when we take into account how the fiber bundle formalism is actually put to use by physicists in representing the physical content of gauge theories, one finds that gauge theories simply are not exhaustively path-relative. In short, even if we assume that the PRI is valid, considerations about mathematical path-relativity in the fiber bundle formalism do not show that gauge theories are exhaustively path-relative as Maudlin suggests. Rather, such mathematical considerations at most establish only this: *If* the physical significance of the fiber bundle formalism is borne exclusively by its mathematically path-relative features, and *if* mathematical path-relativity implies spatiotemporal path-relativity, and *if* gauge theories provide a complete description of fundamental reality, *then* all fundamental physical properties are spatiotemporally path-relative according to gauge theories. But this is a long way from a wholesale revision. Indeed, Lewis and Armstrong can happily grant that such a conditional conclusion follows from Maudlin's argument without having any cause for concern.

To sum up: The claim that gauge theories' use of the fiber bundle formalism calls for a wholesale revision of the Lewis–Armstrong picture hinges on the claim that gauge theories are exhaustively path-relative. But without considering how the fiber bundle formalism

[75] Maudlin (2007), 101. My emphasis.

is actually used by physicists to represent the physical ontologies of gauge theories, mathematical considerations fail to provide reason for thinking that such theories are exhaustively path-relative. The trouble with Maudlin's argument is that it is an explicit attempt to extract the ontology of gauge theories from mathematical features of the fiber bundle formalism. But as we have just seen, it is only by having an antecedent understanding of the ontology of such theories and how the formalism is used to represent it that one can discern which mathematical features are physically significant and which are not.

4 THE CHALLENGE OF INTERPRETATION

Despite my criticisms, I am deeply sympathetic to much of Maudlin's outlook. I wholeheartedly agree that fundamental physics should inform metaphysics. I also agree to some extent that physics paints a revisionary picture of the world. Indeed, I even agree that gauge theories provide evidence that *some* fundamental properties like quark color are spatiotemporally path-relative. My criticism of Maudlin's argument, therefore, has not been directed at this broad outlook, nor towards the claim that physics calls for revisions to contemporary metaphysics, nor even towards his contention that gauge theories posit some path-relative ontology. Rather, my criticisms have been directed specifically at the overtly mathematical character of his argument. My point has been a modest one: The mere fact that gauge theories employ the fiber bundle formalism does not, on its own, lead to the conclusion that gauge theories call for revisions to metaphysics. So although I've suggested that Maudlin's argument ultimately does not succeed, it seems to me that the general spirit it embodies—that of taking science seriously—is spot on. Let us therefore conclude in this spirit by drawing several lessons for future pursuits of scientifically informed metaphysics. In particular, let us consider how Maudlin's argument sheds light on the question: *How is a scientific theory's mathematical formalism related to its ontology?*

In the first place, we have seen that the relationship between the mathematics and metaphysics of scientific theories is far less clear-cut than commonly believed. For one thing, radically different mathematical formalisms can be employed to represent the same physical ontology. For example, as we saw in section 2, the tensor and fiber bundle formulations of EM have the same physical content despite a multitude of mathematical differences. So mathematical differences are not necessarily indicative of physical differences. Moreover, as we saw in section 3, there is no straightforward connection between the *presence or absence of mathematical features* in a theory's formalism on the one hand, and the *existence or nonexistence of entities its physical ontology* on the other. These complications point to a difficult but important question about the interpretation of science: How are we to understand what scientific theories say about the world given their heavy use of mathematics on the one

hand, and the lack of a straightforward connection between a theory's mathematics and its metaphysics on the other hand?

Second, our discussion suggests that the key to addressing this question lies in considering how scientists *actually use* a theory's mathematical formalism. While most philosophers recognize that actual usage must be factored into interpretation, many have failed to fully appreciate its significance. Above, I said that at least two questions must be addressed when interpreting any scientific theory's mathematical formalism: (i) Which formal features are representationally significant? (ii) What is the representational significance of these features? It is largely uncontroversial that facts about how a theory's formalism is used are required to answer the latter question. If, say, I tell you that I'm representing something with the real numbers, you simply do not have enough information to determine π's representational significance—after all, I could be using the real numbers to represent temperatures, or masses, or distances to the moon, or any number of things. Philosophers have therefore largely accepted that one must consider how a mathematical formalism is actually used if one is to answer the latter question. But many philosophers have taken a markedly different attitude towards the first question and held, if only by implication, that *which* features of a theory's formalism are representationally significant depends only on the formalism's structure. This attitude is encouraged in part by the belief that representation involves two sharply delineated factors, *form* and *content*, and that while the content of a representation is highly variable, its form is not. For example, it is thought that while the content of the equation '$F = ma$' varies with usage—in some contexts F represents the force of gravity, in others the force of friction—its form does not. Thus, it is commonly felt that *which* features of a theory's formalism are representationally significant is independent of how that formalism is actually used.

But our discussion suggests that this outlook is fundamentally misguided. Examples like the electric potential illustrate that facts about usage matter not only from the perspective of *what* a mathematical formalism represents, but moreover from the perspective of *which* features of the formalism are representationally significant. The essential point, therefore, is that whether some mathematical feature is representationally significant depends, at least in part, upon how the formalism *is used*. Consequently, whether some feature of a mathematical formalism is representationally significant is not itself a wholly mathematical matter. And this suggests, more generally, that there is no clean divide between the form and content of mathematical representations in science. Or better: there is no such divide independent of how a formalism is put to use. While some mathematical features certainly have representational significance and others clearly do not, such demarcation emerges only against a backdrop of usage.

So here, then, is what I see as the main lesson to take away from our discussion of Maudlin's argument. Investigating a scientific theory's mathematical formalism is surely important to understanding its physical content, and hence its metaphysical implications. But the success of such an investigation requires us to draw upon a much broader body of information—one that includes, amongst other things, facts about scientists' attitudes and practices involving the theory's formalism.

3

THE EXEGETICAL IMPORT OF REPRESENTATIONAL ARTIFACTS

THE BASIC QUESTION of scientific exegesis, *What does science say about the world?*, is not hard to motivate. After all, if science is one of our best sources of knowledge, then surely it is imperative to understand what science says. But just what is it to interpret science? This chapter examines several answers to this question. My contention will be that ***The Problem of Representational Artifacts***, as I shall refer to it, not only shows extant accounts of scientific exegesis to be inadequate, but more importantly, it reveals aspects of scientific exegesis that have largely been overlooked.

1 THE PROBLEM OF REPRESENTATIONAL ARTIFACTS

To begin, let's consider some examples of representational artifacts and the problem they present.

Blackholes and Coordinate Singularities

When Einstein published his theory of general relativity in 1916, it broke ground on a number of fronts: it treated gravity as a local phenomena, thereby doing away with the action at a distance that afflicted Newton's theory; it eschewed an *a priori*, fixed notion of space and time; and so on. General relativity was also ground breaking from a mathematical perspective: it was one of the first physical theories to make extensive and elegant use of differential geometry. This is reflected in the fact that a model of a general relativity is often (though not always) specified as an ordered pair of the form $\langle M, g_{\mu\nu} \rangle$, where M is a mathematical space called a 'smooth manifold' and $g_{\mu\nu}$ a 'metric' defined on M.[1]

A central task of interpreting general relativity, therefore, is to articulate the physical significance of M and $g_{\mu\nu}$. In particular, because Einstein's theory is first and foremost a

[1] More precisely, $g_{\mu\nu}$ is a rank-2 tensor, meaning that $g_{\mu\nu} : TM \otimes TM \to \mathbb{R}$.

theory of gravity, the exegetical task is to determine how M and $g_{\mu\nu}$ represent facts about gravity. And gravity, as Einstein understands it, is closely tied to geometry: roughly, the curvature of M at a point ℓ is proportional to the gravitational force at the spacetime location represented by ℓ. Thus, when a manifold's curvature 'blows up' and becomes infinitely great, this would *seem* to represent a spacetime region where extreme gravitational forces are at play — that is, it would *seem* to represent a black hole.

Unfortunately, it's not that simple. The trouble can be illustrated by considering what is known as ***The Schwarzschild Solution***: an important spacetime model discovered by Karl Schwarzschild just months after the publication of Einstein's theory. The salient feature of Schwarzschild's model is that its manifold contains *two* mathematical singularities. This raises the interpretive question of what spacetime would be like if it were accurately represented by Schwarzschild's model. So if we were to interpret scientific theories by taking them at face value, then we would be led to the following conclusion: that to be accurately represented by Schwarzschild's model, spacetime would have to have two regions where extreme gravitational forces were at play, that is, *two black holes*.

But this is too quick. For here's the thing: it is not merely that there are some reasons to doubt that the Schwarzschild solution actually says that there are two black holes, but rather that there are *serious reasons internal to science* to believe that this is not what it says. Namely, while the singularities may represent black holes, they may also be *mere mathematical artifacts* of Schwarzschild's coordinate system. As one general relativity textbook puts it, "This singular behavior of the components could be due to either (i) a breakdown of the coordinates…or (ii) a true singularity of the spacetime structure."[2] So here we have a serious reason for doubting the naïve interpretation which holds that the Schwarzschild solution posits the existence of two black holes.

But suppose we throw exegetical caution to the wind and let ourselves be swept along by the satisfyingly simple slogan, *Take the claims of science at face value*. Then what? Well, as it turns out, physicists now agree that *only one of the two* mathematical singularities in Schwarzschild's model is physically significant. The other singularity—what came to be known as 'the coordinate singularity'—is a *representational artifact* of Schwarzschild's model. It is no more representationally significant than the color of font in which the mathematics was typeset. Consider what this means. If we had taken Schwarzschild's model to describe *two* black holes, we would have taken it to posit the existence of something that, in fact, it does not. The Schwarzschild solution simply does *not* say that there are two black holes; the naïve interpretation just gets it wrong. Conversely, if we had consigned both mathematical singularities to the status of mere artifact, then we would have similarly misunderstood

[2] Wald (1984), 124.

the content of Schwarzschild's model, though this time by taking it to describe less than it actually does.

The Schwarzschild solution, in addition to providing an example of The Problem of Representational Artifacts, also illustrates The Problem's importance. Above, I said that scientists *now* agree that only one of the singularities is physically significant. But this was not always the case. Indeed, it wasn't definitively shown until 1960 — a full 44 years after Schwarzschild's discovery — that one, and only one, of the mathematical singularities represents a real spacetime black hole.[3] Prior to this, there was an open interpretative question about the content of the Schwarzschild solution. There are two things worth noting about this. First, this sits poorly with a popular conception of representational artifacts that holds that they are merely features that *were not intended to be representationally significant*. But this simply is not what happened in the case of the Schwarzschild solution — it was not as if Schwarzschild simply sat down and decided that one singularity would be significant and the other would not. This means that something beyond the intentions of scientists must be at work in determining which features are representationally significant and which are not. Second, it's important to appreciate the fact that the exegetical obstacle presented by representational artifacts is *not* some worry dreamed up in the philosophical armchair. Quite to the contrary, The Problem of Representational Artifacts has, on numerous occasions, slowed the progress of actual science. The Schwarzschild solution is one such case, but general relativity contains a treasure trove of other, better examples of how misunderstandings about the nature of mathematical artifacts can stand in the way of scientific progress. Take Einstein, for example. After publishing his special theory of relativity in 1905,[4] it took him nearly a decade to extend the account to handle gravitational phenomena. The hangup? Representational artifacts. In Einstein's own words: "Why were another seven years required for the construction of the general theory of relativity? The main reason lies in the fact that it is not so easy to free oneself from the idea that coordinates must have an immediate metrical meaning."[5]

Wolves, Rabbits, and Real Numbers

The ***Lotka-Volterra model*** (LVM), or what is known more popularly as the 'predator-prey model', is a mathematical model used in ecology for modeling population dynamics—

[3] Prior to 1960 a number of physicists had suspected that there was an important difference between the two singularities, for instance see Lemaître (1933). However, it wasn't until 1960 when Martin Kruskal produced a single coordinate system that covered the entire Schwarzschild spacetime that it was definitively shown that one of the singularities was a mathematical artifact. See Kruskal (1960).

[4] Einstein (1905).

[5] Einstein (1949), 67.

e.g., modeling correlations between the growth and decline of wolf and rabbit populations.[6] Let's consider how The Problem of Representational Artifacts arises when interpreting the LVM.

If all you knew about the LVM was that it was a model of wolf and rabbit populations, then it's understandable how it might sound plausible that we should take the LVM at face value — more rabbits means more wolves, more wolves means fewer rabbits, fewer rabbits means fewer wolves, and so on. What's the trouble with taking something like that at face value? But here's the thing: The talk of wolves and rabbits lends to an impression of concreteness that is nowhere to be found when we actually look at the LVM. Instead, what we find is a system of differential equations:

$$\frac{dx}{dt} = rx - axy, \tag{LVM1}$$

$$\frac{dy}{dt} = \varepsilon axy - my, \tag{LVM2}$$

You have to admit, it's a bit hard to take these equations at face value. After all, what would it even *mean* to take (LVM1) and (LVM2) at face value? So it's hard to deny that some amount of exegetical work is needed to understand the LVM.

Minimally, we need to know what the LVM's variables represent:

LVM Variable	Representational Significance
x	prey population
y	predator population
r	population growth rate
m	population morality rate
a	predator capture efficiency
ε	conversion efficiency

But let's suppose that we know this. Could we then take the LVM at face value? It's a tempting thought. Once we know what the LVM's variables represent, and assuming we understand the ecologist's basic lexicon, then (LVM1) seems to be nothing more than a compact way of expressing the claim:

(LVM1*) The change in prey population over time is equal to the product of the prey's population with its growth rate, minus the product of the predator population, prey population, and the predator's capture efficiency.

[6] The LVM is discussed in virtually every book on (ordinary) differential equations. Here I am drawing on Logan (2011), 232–236.

While this is beginning to have the appearance of something that could be taken at face value, some exegetical work remains. We need to know what the ecologist means by terms like 'capture efficiency' and 'conversion efficiency'. But what if we also assume that we understand the ecologist's basic lexicon, e.g., that 'capture efficiency' is just a fancy term for the chance that a predator will walk away with dinner on any give encounter? In this case, can we finally take the LVM at face value?

No. And here's the trouble: Even after going through the basic rigamarole of lexical exegesis, the LVM contains representational artifacts that prevent taking it at face value. This happens in a slightly different way than in the case of the Schwarzschild solution. The Schwarzschild solution is a physically reasonable model of general relativity containing a representational artifact. The LVM, however, is a class of models (i.e., the class of solutions to the above equations). And this class of models exhibits at least two different varieties of representational artifacts. In the first place, not all mathematical solutions to the LVM are models of possible ecosystems. For example, in many of the mathematical solutions to the LVM, the variables representing the predator and prey populations take negative values. But of course, no possible ecosystem contains -10 wolves. The LVM contains a second, even more vivid example of representational artifacts. The thing we need to notice is that the predator and prey variables, x and y, are continuous real-valued functions. Let's unpack this. To say that they're real-valued functions is straightforward enough: It simply means that x and y are functions from 'time' (usually represented by the unit interval $[0, 1]$) to the real numbers.[7] The important fact is that they are *continuous* functions, meaning that if they take different values at different times, then they must pass through *every* real number in between those values. For instance, consider a solution to the LVM where the prey population increases from three rabbits at t_1 to five rabbits at t_2. Then the fact that x is a continuous function means that for any real number r between 2 and 5 there must be some time t (between t_1 and t_2) where $x(t) = r$. So here's the punchline: Even with all our exegetical work, if we were to take the LVM at face value then we would be led to misunderstand its content. On the face-value reading, the LVM describes ecosystems with π number of rabbits. But undetached-rabbit-parts notwithstanding, this is clearly impossible. So observe: Not only does the LVM contain physically unreasonable models, its reasonable ones still present exegetical obstacles.

This is not an idle point. After all, surely some scientific theories turn out to be logically inconsistent, to describe physically impossible states of affairs, or to suffer from any other number of defects. And surely the correct interpretation of those theories must recognize the fact that they have such defects. So when we uncover such a defect, there are two ways we might go. Either we can decide that the fault lies not with science but instead with our

[7] I.e., $x(t) : [0, 1] \to \mathbb{R}$, and similarly with $y(t)$.

interpretation, or we can decide that our interpretation is correct and that there is some flaw in the picture that science paints of the world. So one significant trouble with taking science at face value is that, while commonsense might serve to illuminate such a fork when we come to it, surely it doesn't tell us which path to follow. For instance, commonsense might tell us that it's impossible for there to be -10 wolves, but commonsense doesn't tell us whether we have misinterpreted the LVM or if we have instead correctly interpreted a flawed model.

The Problem of Representational Artifacts

The Schwarzschild solution and the LVM are not isolated examples. Quite to the contrary, *every* scientific representation contains representational artifacts. Indeed, any method of communication possesses attendant aspects that are incidental to the information conveyed — the language in which a message is expressed, or the typeface and color of ink with which it is printed, are paradigms of such attendant aspects. And knowing what to take seriously and what to disregard is crucial to communication. That is, linguistic competence presupposes an ability to differentiate between those aspects of language that are integral to the information conveyed and those which are merely incidental. So let's take a step back and consider The Problem of Representational Artifacts's broader significance for scientific exegesis.

The main point is that understanding what science says requires that we first identify which aspects of science are representationally significant and which aspects are not. After all, as we saw with the coordinate singularity, if one mistakenly believes some aspect is significant when it is not, then one will understand science to say something that, in fact, it does not say. And if one disregards some aspect of science that is representationally significant, then one will take science to say less than it actually does. Thus The Problem of Representational Artifacts points to two equally egregious exegetical errors. On the one hand, an interpretation might mistake a representational artifact for a substantive aspect of science. On the other hand, it may err in the opposite direction, by mistakenly understanding some substantive aspect of science to be a mere artifact. So here, then, we have two different exegetical errors, but the effect is the same in both cases: failure to understand what science says about the world.

Having gotten a feel for The Problem of Representational Artifacts, let's turn to considering its implications for extant theories of scientific exegesis.

2 SCIENTIFIC LITERALISM

The most naïve view of scientific exegesis is what might be called ***Scientific Literalism***: the view that interpreting science amounts—big surprise—to taking the claims of science literally. In a moment we'll ask just what this means, but I want to begin by highlighting several aspects of Literalism's general exegetical outlook.

When substantive questions about what science says arise, Literalists tend to feel that such questions should be settled by scientists, not philosophers. There is something to this, of course. After all, if you're unclear about what it means to say "Mitochondria have outer membranes", then shouldn't you consult a biology text? Or even better, why not walk across campus and consult an actual biologist? For if the biologist is the one making the claim, Literalists will argue, shouldn't the biologist be the one to say what it means? So for many Literalists, the primary challenges to understanding science are not exegetical, but metaphysical and epistemological. That is, Literalists tend to feel that the hard, philosophically important questions are not questions about *what science says*. Instead, they hold that the important questions concern whether we are *justified in believing* what science says, whether we should *posit* unobservable entities or only admit into our ontology those entities with empirical consequences, and so forth.

The obvious question for the Literalist is just what it means to take the claims of science literally. That is, what differentiates literal and metaphorical interpretations of science? One popular suggestion looks to the concept of structure. Here's how Bas van Fraassen describes it:

> The decision to rule out all but literal construals of the language of science, rules out those forms of anti-realism known as positivism and instrumentalism. First, on a literal construal, the apparent statements of science really are statements, *capable of being true or false*. Secondly, although a literal construal can elaborate, it *cannot change logical relationships*. … On the positivists' interpretation of science, theoretical terms have meaning only through their connection with the observable. Hence they hold that two theories may in fact *say the same thing although in form they contradict each other.* But two theories which contradict each other in such a way can 'really' be saying the same thing *only if they are not literally construed.* Most specifically, if a theory says that something exists, then a literal construal may elaborate on what that something is, but will not remove the implication of existence.
>
> To insist on a literal construal of the language of science is to rule out the construal of a theory as a metaphor or simile, or as intelligible only after it is 'demythologized' or subjected to some other sort of 'translation' that *does not preserve logical form*. If the theory's statements include 'There are electrons', then the theory says that there are electrons. If in addition they include 'Electrons are not planets', then the theory says, in part, that there are entities other than planets. …

> Not every philosophical position concerning science which insists on a literal construal of the language of science is a realist position. For this insistence relates not at all to our epistemic attitudes toward theories, nor to the aim we pursue in constructing theories, but *only to the correct understanding of what a theory says*. (The fundamentalist theist, the agnostic, and the atheist presumably agree with each other (though not with liberal theologians) in their understanding of the statement that God, or gods, or angels exist.)[8]

So according to this proposal, literal interpretations of science have two key features: they are *truth apt* and they *preserve logical form*.

Before asking what this latter condition comes to, it's worth pausing to reflect on the final paragraph of this quote. What Van Fraassen correctly points out is that exegetical questions are logically prior to those about the epistemology and metaphysics of science. The question of scientific realism, for example, is logically downstream from the exegetical question of what science says: before the realist and antirealist can debate about how much of science to believe, first they have to agree about what science says. Before the metaphysician can argue that one theory should be preferred to another on, say, the grounds of parsimony, she has to know that the first theory in fact posits more than the second—that is, she must first know what the two theories say.

Now focus on the latter condition for being a literal interpretation: preserving logical form. Just what is it for an interpretation to preserve the logical form of a scientific representation? We'll take up this question in its full generality below, but for now let's concentrate on the question in the case of scientific theories and their statements. In these cases, the most common way to understand what it is to preserve logical form is what, in some contemporary philosophical circles, might be called the 'neo-Quinean' approach.[9] Now, Neo-Quineans tend to spill a lot of ink over the *existential quantifier* of formal logic,[10] but it seems to me that the view is better epitomized by a belief in a gratifyingly simple connection between ideology and ontology: roughly, constant terms correspond to individuals, one-place predicates to properties, two-place predicates to binary relations, and so forth. Of course, this is easy enough to flesh out in quantificational terms. Say that the

[8] van Fraassen (1980), 10-11. My emphasis.

[9] *Nota bene*: Contemporary neo-Quineans frequently attribute their views concerning the nature of ontology to the theses (apparently) defended in Quine (1948). But much of the apparent similarity, it seems to me, results from reading Quine (1948) out of its historical context, especially out of the context of Quine's other writings. For while there's no denying that the real Quine extolled the virtues of first-order logic (e.g., see Quine (1960), ch. 5) and frequently advocated the existential quantifier as a criterion of ontological commitment, we must keep sight of the pragmatism that underpins Quine's thought. For instance, unlike neo-Quineans, the historical Quine would deny that the existential quantifier of first-order logic furnishes an *objectively correct* or *ontologically privileged* notion of quantification (e.g., see Quine (1951) or Quine (1969)); quite to the contrary, Quine is clear that he sees the reasons for preferring the existential quantifier stem from pragmatic considerations such as its utility for simplifying theories (e.g., see Quine (1960), 143–147).

[10] For example, see van Inwagen (1998).

purely structural commitment of "Φ" is the statement "PSC(Φ)" generated by existentially generalizing over every non-logical symbol in "Φ". For example, the purely structural commitment of

$$\text{Taller}(\text{Tom}, \text{Mary}) \wedge \text{Happy}(\text{Mary})$$

would be:

$$\exists X \exists Y \exists x \exists y \, (X(x, y) \wedge Y(y)) \, .$$

We could debate (of course) about the precise rules for generating a statement's purely structural commitment: Should, for instance, generalizations over predicate and relation symbols be understood in the familiar set-theoretical way or in terms of plural quantifiers?[11] What should we take as the purely structural commitment of a negative existential statement, e.g., "Pegasus doesn't exist"?[12] And so on. While Literalists may disagree over such details, they are united in thinking that a literal interpretation of a statement "Φ" must entail "Φ" 's purely structural commitment. That is:

(LIT) $\mathscr{I}(\Phi)$ is a literal interpretation of "Φ" only if $\mathscr{I}(\Phi)$ entails PSC(Φ).

With this in mind, the basic exegetical commitment of Scientific Literalism may then be understood as an adherence to the schema:

(DS) If a theory T's statements include "Φ", then T literally says that Φ,

where this the consequent of (DS) conforms with (LIT).[13]

Now, Literalism is clearly an inadequate theory of scientific exegesis. But it's important to see why. For what I'm going to suggest is that it is inadequate twice over: it fails to provide an adequate account both of the *epistemological* aspects of scientific exegesis as well as its *metaphysical* aspects. In other words, it fails to provide any serviceable answer to the epistemological question, *How should we determine what science says?* But more fundamentally, it misunderstands the nature of scientific exegesis.

2.1 THE EPISTEMOLOGICAL LIMITATIONS OF LITERALISM

Let's begin with Scientific Literalism's epistemological inadequacies, for they tend to be more blatant than its metaphysical failings.

[11] Cf., Boolos (1984).

[12] Cf., Quine (1948).

[13] This leaves several important questions unresolved. One is whether (DS) should be strengthened to a biconditional—that is, is "Φ" being included in T's statements not only sufficient but also necessary for T to say that Φ? Another is whether (LIT) provides not only a necessary for demarcating the literal from the metaphorical, but also a sufficient condition. Both of these are interesting questions that I'll return to below. At present, however, simply take note of the fact that Literalism comes in varying strengths depending on whether (DS) and (LIT) are strengthened to biconditionals.

A first epistemological difficulty is that, owing to its adherence to (DS), Literalism allows for completely uninformative interpretations of science. For instance, given a sentence like "There are electrons", Literalism approves of the interpretation:

(1) If T's statements include "There are electrons", then T says that there are electrons.

This is obviously unsatisfactory. By way of illustration, consider the following interpretation of Lewis Caroll's poem *Jabberywocky*:

(2) If *Jabberywocky*'s statements include "All mimsy were the borogoves" then the *Jabberywocky* says that all mimsy were the borogoves.[14]

The problem here is that if you don't *already know* what "All mimsy were the borogoves" means then you're *not going to learn* by being told that it means all mimsy were the borogoves. And similarly, Literalism fails to provide a *genuinely informative account* of what the world would have to be like to be as science says it is.

It's easy to see Literalism's troubles here as less severe than they really are. The temptation is to diagnose Literalism's troubles as arising solely from the fact that it performs poorly when confronted with statements, theories, etc., containing *undefined terms*. Take "All mimsy were the borogoves" as an example. The reason this fails to express a determinate, nontrivial claim is obviously that, as they stand, the terms 'mimsy' and 'borogoves' have about the same significance as scribbles. But there is an easy solution, of course: just define them! And once our definitions are secured, the sentence is no longer a mere string of symbols, but (we may imagine) expresses a determinate claim. So it's tempting to think that defining our terms is what makes all the difference between full-blooded scientific representations and mere symbols. Thus, we should understand the Literalist's basic exegetical commitment not as to (DS), but instead to:

(DS*) If a theory T's statements include "$\Phi(\ldots\tau\ldots)$" and 'τ' means $\tau^\star$, then T literally says that $\Phi(\ldots\tau^\star\ldots)$,

where $\tau^\star$ is some antecedently understood expression.[15]

But a serious difficulty arises when we turn to the epistemological question of *how to determine* what science says. Literalism maintains that this requires nothing more than determining which statements science affirms. If we want to determine, say, whether a theory T says that there are electrons, all we have to do is see if T's statements include "There are

[14] Carroll et al. (1990), 195–196.

[15] Of course, (DS*) raises at least two important, general metaexegetical questions: *What does it take to define an unknown scientific term (expression, etc.)?* And: *Given an unknown term (expression, etc.) τ, is there, in general, a best way to define τ?* I'll return to these questions later on.

electrons". Yet obviously this isn't an answer—at least not really. For just how are we supposed to determine if T's statements include "There are electrons"? Literalism doesn't say, and this is clearly a problem. So to be worth its salt, any theory of scientific exegesis had better answer the question: *How should we determine which statements science affirms?*

I can imagine a number of proposals that a Literalist might float, but let me skip ahead to two plausible suggestions frequently found in the philosophy of science: *Just ask the scientists!* And: *Simply consult science textbooks, journal articles, etc.*[16] These answers can seem so commonsensical that they hardly warrant debate. And yet, neither is satisfactory. For one thing, at least as a contingent matter of fact, scientists are quite often unaware of the full range of statements comprising their own theories. Maxwell, for example, initially did not know that his theory said that light was a form of electromagnetic radiation; indeed, it was only later as he reflected upon the behavior of a constant in his equations that this realization came to him. The Literalist will likely chalk this up to a failure of logical omniscience—a failure that occurs (roughly) when a set of statements Γ entails a statement φ but, owing to cognitive limitations, one fails to see that φ is a consequence of Γ. In such cases, while the scientist understands what her terms *mean*, what her theory *says*, and so forth, she fails to fully trace out their consequences. So the key thing about failures of logical omniscience is that they are *epistemic* failures, not *exegetical* ones. And if this is so, the Literalist may contend that scientists' ignorance about their own theories is no mark against Literalism.

The Literalist is certainly correct that failures of logical omniscience do happen in both the physical sciences as well as in logic and mathematics.[17] (After all, if such failures didn't occur, mathematicians would be out of jobs!) Yet even if many instances of ignorance are attributable to such epistemic failings, the question for the Literalist is whether *all* cases can be so understood. Are there, in other words, instances where scientists' ignorance of their own theories' content is the result of an exegetical failure — that is, a failure to understand the meaning of their own terms, the physical significance of their own equations, etc.? Such failures may seem hard to fathom. Isn't this akin to suggesting that a speaker may fail to know what he means by his own assertion? And while there are ways that this can happen—if, say, the speaker engages in mere phonological or orthographic mimicry, or if by 'means' we mean something in lines with an externalist theory of meaning, etc.—it is hard to see how an ordinary speaker in ordinary circumstances can fail to understand, in an ordinary sense, the meaning of his own assertions. And likewise in the scientific case.

But here's the thing: such failures *do* occur in science. Allow me to highlight two. First, as we've already seen, general relativity is a veritable treasure trove of examples. As I noted,

[16] For example, in discussing what classical mechanics says about a particular model known as 'the harmonic oscillator', Ronald Giere writes: "I propose we regard the simple harmonic oscillator and the like as abstract entities *having all and only the properties ascribed to them in the standard texts*." (Giere 1988, 78).

[17] See Garber (1983).

when Einstein was wrestling with what would eventually become his general theory of relativity, his efforts were hampered by confusions about the physical significance of coordinate systems. In essence, Einstein was hung up on whether differences in how locations were labeled amounted to physically significant differences. Worse, after he finally realized that such differences were physically *in*significant, he came to believe that that this fact, what came to be called 'general covariance', was in of itself deeply telling about the physical significance of general relativity. It was quickly pointed out, however, that general covariance had no physical significance but was, instead, merely a reflection of a physicist's mathematical cleverness at relabeling spacetime coordinates.[18] Einstein had misunderstood the physical significance of coordinates twice — and in both cases, his misunderstanding was exegetical, not epistemic, in origin. A second example comes from Faraday's theory of electromagnetism. On Faraday's interpretation, 'line of forces', as he called them, were the central, physically significant element of his theory. However, as Maxwell began to study Faraday's work, he came to the realization that lines of force are, *pace* Faraday's interpretation, *not* real physical entities. Lines of force are, rather, a conceptual crutch, a helpful heuristic, but they are nothing more. They are, in Maxwell's words, "a purely imaginary substance".[19] Which is to say, Faraday misunderstood the physical significance of a central part of his own theory — again, not for epistemic reasons, but for exegetical ones.

Neither example is entirely clearcut, of course — we could argue, for instance, about whether general covariance actually is physically significant.[20] But the point here is simply that *there is something to argue about*: in both examples, there are substantive *exegetical* questions about the physical significance of scientific representations. Further, in each case, the scientists involved were arguably mistaken, at least partly, about the content of their own theories. Further still, their confusion arose from conceptual challenges surrounding exegetical notions like representational equivalence, the physical content of a scientific representation, and, of course, The Problem of Representational Artifacts.

So one reason that scientific exegesis isn't just as simple as asking the scientists or consulting the scientific literature is that scientists can be ignorant of what their own theories say, and not just because they have failed to fully work out their theories' consequences, but more fundamentally because there is a lack of clarity surrounding the nature of scientific exegesis. Consequently, there is a good metaexegetical question, *How should we determine what science says?*, that can't be answered by merely deferring to scientists or the extant scientific literature. And this means that, at least from from an epistemological point of view, Scientific Literalism is a failure.

[18] This objection was raised by Kretschmann (1917).
[19] Maxwell (1890), 160.
[20] For instance, see Norton (2003).

2.2 THE METAPHYSICAL LIMITATIONS OF LITERALISM

Literalism's epistemological shortcomings are symptoms of deeper, metaphysical problems. For the unsatisfying thing about Literalism from a metaphysical point-of-view is that it fails to shed any light on the nature of scientific exegesis.

Think of it this way: Even if asking the scientists and consulting the textbooks produced an extensionally correct understanding of what science says—which, as we have just seen, there is good reason to think it almost certainly does not—it would reveal nothing about concepts like representational equivalence and physical significance, nor about what separates full-blooded scientific representations from arbitrary collections of signs and symbols, nor about what it is for something to be a representational artifact, nor about any number of other issues surrounding the nature of scientific exegesis. Imagine that Literalist posed the question of what some theory says to a scientist. What should the Literalist do if the scientist doesn't understand the question? Or better, imagine yourself in the scientist's shoes. You want to figure out what your own theory says. Soliloquy notwithstanding, you can't just ask the scientists. Nor can you just consult the textbooks, journal articles, and so on—after all, you want to know what your theory says so that you can write the textbooks, journal articles, etc. So what are you to do? How should you go about determining what statements your theory includes? More fundamentally, what are you even trying to do?

Now, it can be hard to hear what it is that this latter question is asking, but in a way, that's precisely the point. There is a good question here, but owing to Literalism's limited conception of interpretation, it's difficult to even state.

A better appreciation of this conception's limitations can be gained by considered ***mixed mathematico-physical statements***, statements like:

(3) The empty set is as fluffy as a poodle.

(4) The real numbers are part of a well-balanced diet.

Statements like these are not just false, they're obviously false on conceptual grounds: it's a gross mistake to ascribe physical properties to mathematical entities. Conversely, it is no less erroneous to ascribe mathematical properties and relations to physical objects, e.g.;

(5) Joesph Conrad is a prime number.

(6) Trump is divisible by Clinton.

While few would deny the obvious incoherence of (3)–(6), not all mixed statements are manifestly meaningless mélanges of mathematics and the mundane. Quite to the contrary, many mixed statements are less obviously problematic. For instance, you don't have to look hard to find examples like:

(7) Spacetime is a differentiable manifold.

(8) Momentum is a self-adjoint operator.

Observe that these statements are every bit as confused as (3)–(6) if they are interpreted literally. After all, spacetime and momentum are elements of physical reality; differentiable manifolds and self-adjoint operators are mathematical entities. This strongly suggests that the correct interpretation of (7) and (8) is not the literal one. But perhaps the Literalist can accommodate this fact in a fashion similar to how she dealt with unknown terms. For instance, (7) and (8) would become:

(7*) Spacetime is ψ,

(8*) Momentum is χ,

where ψ and χ are physically significant expressions.

While this strategy may enjoy some success, it fails to furnish a general solution to the problem of mixed mathematics. The trouble is that for the vast bulk of mixed statements, interpretation requires far more than mere substitution. Take (7) as an example. If we're looking for a statement that can be understood literally, then the obvious candidate is not (7*), but instead:

(7**) Spacetime *is represented by* a differentiable manifold.

However, the Literalist cannot accept (7**) as an interpretation of (7). For one thing, (7**) does *not* result from merely replacing the mixed idioms in (7) with physically significant expressions. In other words, (7**) does *not* entail the purely structural commitment of (7) and so fails to conform with (LIT). So it can happen that the correct interpretation of an expression "Φ" does *not* entail $\textsc{psc}(\Phi)$.

There is deeper reason that the Literalist cannot accept (7**) as an interpretation of (7). Namely, any victory that the 'represented by' idiom secures for the Literalist over the problem of mixed mathematics is, at best, a pyrrhic one: absent an account of what it is for a differentiable manifold to represent spacetime (7**) is entirely vacuous, (as I shall argue in detail below). To avoid vacuity, the Literalist's interpretation of (7) would have to be (7**) *in conjunction with an account of the representational role played by differentiable manifolds within spacetime theories*. And then…suffice it to say, we're a long way from just taking science at face value.

So look at the problem here this way: If scientific exegesis consisted in nothing more than definition, substitution, and the like, then it would be easy to think that there wasn't a whole lot to interpreting science. But if the correct interpretation of mixed statements is one that involves the idiom of representation, then there must be more to the nature of scientific exegesis.

3 RAMSIFICATIONISM

Let's turn a second, more plausible account of scientific exegesis: ***Ramsificationism***. Unlike Literalism, which was at best a caricature, Ramsificationism has been endorsed in varying forms by philosophers ranging from logical positivists the likes of Rudolph Carnap, to metaphysicians such as David Lewis and Frank Jackson.[21] It's not hard to see why. Ramsification promises to resolve many of Literalism's problems in a simple way: roughly, by appealing to the functional roles of unknown aspects of science to pin down their interpretation.

To understand Ramsificationism, a good place to start is with the semantic empiricists (Locke, Hume, *et al.*) who held that term, concept, etc., are meaningful *only if* they refer to immediate elements of experience (e.g., 'yellow') or are defined in terms of such elements (e.g., 'golden mountain').[22] The trouble is that science contains lots of terms that are neither. For instance, consider the terms 'force' and 'mass'. Clearly, these are *not* immediate elements of experience. But it also seems that neither can be defined in terms of such empirical elements. For in light of Newton's second law—*force equals mass times acceleration*—defining 'force' requires that we say something about what 'mass' means, and vice versa. So it appears that neither term can be defined without invoking the other, non-empirically definable term. And if this is so, then it looks like the semantic empiricist has been forced into the unenviable position of saying that, e.g., Newton's second law is *not* meaningful. What to do?

Enter Ramsificationism. Its crucial innovation was to liberalize what counts as a definition. Up to this point we have been assuming that a definition's *definiendum* cannot appear amongst its *definiens*—in other words, what's being defined can't be part of its own definition. When combined with the semantic empiricist's requirement that concepts be definable in empirical terms, we end up with a fairly strict conception of definition: that the definition of a constant symbol 'c' or of an n-place relation 'R' must take the form

$$\forall x\big(x = c \leftrightarrow \eta(x)\big),$$
$$\forall x_1 \ldots \forall x_n\big(R(x_1, \ldots, x_n) \leftrightarrow \epsilon(x_1, \ldots, x_n)\big),$$

where 'η' and 'ϵ' contain only empirically meaningful terms. To relax matters, the Ramsificationist defines a theory T's unobservable terms through, roughly, the functional roles

[21] As examples, see Carnap (1995); Lewis (1970) and Lewis (2009); and Jackson (1998).

[22] One of the main criticisms of the empiricist's appeal to Ramsificationism was that hinged on the untenable division between observable and unobservable terms (E.g., Maxwell (1962).) But the basic insight of Ramsificationism does not depend upon this distinction. Quite to the contrary, all Ramsification procedure requires is a distinction between the antecedently understood elements of a theory and those elements which stand in need of interpretation (Lewis (1974).) Given this undemanding distinction, the Ramsification procedure then allows us to interpret these elements by the role they play within the theory as a whole. So not only is there no need to worry that Ramsificationism rests on shaky empiricist assumptions.

they play within T. The Ramsification procedure involves five basic steps.[23] First, take all the statements T affirms and write them as one long conjunction, $\mathfrak{P}_T$, called T's ***postulate***. Second, divide $\mathfrak{P}_T$'s vocabulary into two classes: observable and unobservable terms. Third, systematically replace all of $\mathfrak{P}_T$'s unobservable terms $\tau_1, \ldots, \tau_n$ with free variables $x_1, \ldots, x_n$. Fourth, existentially generalize over these free variables resulting in what is known as T's ***Ramsey sentence***:

$$\exists x_1 \ldots \exists x_n \mathfrak{P}_T(x_1, \ldots, x_n).$$

The Ramsificationist views this as the interpretation of T. But what of the interpretation of the individual unobservable terms that appeared in T? The final step, then, is to define these terms by what is known as T's ***Carnap sentence***:[24]

$$\exists! x_1 \ldots \exists! x_n \mathfrak{P}_T(x_1, \ldots, x_n) \rightarrow \mathfrak{P}_T(\tau_1, \ldots, \tau_n).$$

And if T's postulate is uniquely realized, then T's Carnap sentence furnishes an interpretation T's unobservable terms: 'τ_1' refers to whatever realizes x_1, etc.

Ramsificationism has a number of things going for it. For one thing, the Ramsification procedure just is, in point of fact, how some scientists interpret parts of their theories. In electromagnetism, for example, the Lorentz force law

$$\frac{d\mathbf{p}}{dt} \;=\; q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

is sometimes said *to provide a physical definition of the mathematical symbols* **E** *and* **B**.[25,26] When physicists say this, they more or less have in mind the Ramsification procedure. More importantly, from the point of view of understanding the nature of scientific exegesis, Ramsificationism improves upon Literalism in two crucial respects. First, by relaxing the requirements on definitions, Ramsificationism ostensibly does better than Literalism when it

[23] See Carnap (1995), chapter 26, for a detailed overview of the Ramsification procedure.

[24] This is term was introduced by Lewis; see Lewis (1970).

[25] For instance, see Misner et al. (1973), 72, and Jackson (1999), 777–778.

[26] It's important to note that in saying that the Lorentz force law physically defines **E** and **B**, most of the terms and relations in the statement of the law are antecedently understood: **p** is the momentum of a test particle, $d\mathbf{p}/dt$ is the rate of change of momentum over time, q is the charge of the test particle, **v** is the test particle's velocity, and the mathematical operation, $\times$ or 'the cross product', indicates the direction perpendicular to both the particle's velocity and whatever direction **B** points in. Given this, the Lorentz force law could be clumsily rewritten as:

> The rate of change of a test particle's momentum with respect to time is equal to its charge times (**E** plus a vector quantity whose magnitude and direction is proportional with, and mutually perpendicular to, the particle's velocity and **B**).

Replacing **E** and **B** with free variables we have:

> The rate of change of a test particle's momentum with respect to time is equal to its charge times (x plus a vector quantity whose magnitude and direction is proportional with, and mutually perpendicular to, the particle's velocity and y).

And then whatever realizes x and y is called 'the electric field' and 'the magnetic field'.

comes to the question of what differentiates arbitrary strings of symbols from full-blooded scientific representations. Second, unlike Literalism, Ramsificationism appears to make room for genuine instances of scientific novelty. The trouble, recall, was that if τ is explicitly defined in terms of some antecedently understood concepts $c_1, \ldots, c_m$, then τ is not really a genuinely new concept but only an abbreviation for some combination of $c_1, \ldots, c_m$. Thus, if the only notion of definition one recognizes is that of an explicit definition, as the Literalist does, then one faces a dilemma. When a new scientific term τ is introduced, either τ is explicitly defined or it is not. If it is, then τ adds nothing genuinely new to the conceptual resources of science. If it is not defined, however, then τ is no more than an arbitrary symbol which lacks physical significance. In the former case, The Problem of Scientific Novelty rears its head; in the latter, The Problem of Full-Bloodedness. While Literalism offers no escape from the dilemma's horns, the promise of Ramsificationism is that in it we find a *tertium quid*: definition via the Ramsification procedure simultaneously accounts for the difference between full-blooded scientific representations and arbitrary symbols but while allowing for genuine scientific novelty.

3.1 DIFFICULTIES WITH RAMSIFICATIONISM

Although Ramsification is a step forward from Literalism, it is still far from an adequate theory of scientific exegesis. One well-known difficulty with Ramsification comes from a number of so-called 'definability theorems' like Beth's theorem and Svenious's theorem, to cite two of the most well-known. Given some relatively minimal assumptions, these theorems essentially dissolve the difference between explicit definability and definability via the Ramsification procedure. That is, they show that if τ can be defined via the Ramsification procedure, then τ can also be explicitly defined. But the entirety of Ramsification's promise as a simultaneous solution to The Problems of Full-Bloodedness and Scientific Novelty was predicated on the idea that it could define terms while avoiding explicit definitions. A second well-known and much-discussed objection is that Ramsification has a tendency to trivialize theories by reducing them to mere numerical claims.[27] That is, the danger is that if the Ramsey sentence of a theory is true in some world W, then it will be true in every other world that contains the same number of entities as W.

Rather than getting bogged down with these familiar problems, however, I want to focus on two new difficulties with Ramsificationism that shed light on the nature of scientific exegesis. The first is that, like Literalism, Ramsificationism is predicated upon an excessively narrow picture of scientific exegesis. But as we shall see, Ramsificationism is exces-

[27] The problem was first explicitly noted in Newman (1928)'s review of Russell (1927). Newman's characterization of the problem, as well as his proposal for resolving it, prefigured the basic gambit of everything from Carnap (1969) to Lewis (1984). Some particularly helpful recent discussions of the problem include Demopoulos and Friedman (1985), Ketland (2004), and Ben-Menahem (2006), ch. 4.

sively narrow in subtly different ways than Literalism. The second difficulty is that Ramsificationism leads to demonstrably incorrect interpretations of actual scientific theories. That is, it just gets the exegetical facts wrong.

3.2 RAMSIFICATIONISM'S LIMITED EXEGETICAL OUTLOOK

Ramsificationism's exegetical outlook has at least three major limitations that illuminate the nature of scientific exegesis.

Limitation 1: Representational Success. A first, well-known limitation of Ramsificationism is that it rests on some significant assumptions about the success of what is being interpreted. It's well-known, for instance, that Ramsification's success at interpreting a theory T requires that T's postulate have a unique realizer.Now, a great deal has been said about the problem of multiple realization and I don't have time to trudge through all the proposals for solving it, but let me mention two of the most vaunted suggestions. The first suggestion is that the threat posed by the problem of multiple realization is inversely proportional to the size of a theory's postulate. For as the size a theory's postulate increases, so too do the requirements for realizing it. And if it becomes large enough, we may hope that it will exclude all but one possible realizer. For example, van Fraassen writes:

> This suggests that the intention, of which sorts of phenomena are to be embedded in what kinds of empirical substructures, be made part of the theory. *I do not think that is necessary. Unintended realizations disappear when we look at larger observable parts of the world;* say, the optics and mechanics of moving light sources together. If for a little while some fairly weak theory is empirically adequate, but in a way its advocates are not in a position to notice, that seems hardly an important or frequent enough occurrence to guard against by more complex definitions.[28]

A second, related suggestion is to cut down on possible realizers of a theory by appealing to the notion of naturalness. Natural properties are properties like *redness* and *roundness*, properties that explain facts about objective resemblance. They stand in contrast to gerrymandered properties like *grue* and *bleen* that fail to carve nature at its joints, so to speak. So if you believe in natural properties and you think that science is in the business of discovering objective features of the world, then it's only natural to hold that science is in the business of discovering natural properties and describing how they are distributed throughout the world. And if this is so, it seems you have an easy way of eliminating the possibility of multiple realization: namely, insist that in science the only admissible realizers are natural properties. Now, as I said, much has been written about all this and why neither proposal resolves the problem of multiple realization—but I don't have space to dive into any of this

[28] van Fraassen (1980), 66. My emphasis.

here. So, for the sake of argument, let's just grant that if a theory T has a large enough postulate and that if we insist it can only be realized by natural properties, then T will be uniquely realized.

But what exactly does this buy the Ramsificationist? Not much. For one thing, requiring that realizers be natural properties is far too restrictive. After all, much of science does *not* purport to describe the most basic, fundamental elements of reality; e.g., disciplines like ecology and geology describe relatively high-level phenomena. So restricting to natural properties would, at best, allow Ramsificationism to eliminate multiple realizations in relatively narrow number of fields like fundamental physics. For another thing, even if this solves the problem of multiple realizations for a few grand-unified theories, what about the rest of science? What about theories with short postulates or that describe relatively small domains of phenomena? The Ramsificationist may therefore wish to jettison the uniqueness requirement and settle for some weaker conception of definition. One potential way of doing this is reminiscent of supervaluationism: namely, allow for indeterminate definitions so long as the indeterminacy is irrelevant to the truth of claims involving of the term.[29] Whether or not this is so is an interesting question, but not one I will pursue here.So let us set the problem of multiple realizations aside.

Where Ramsificationism really runs into trouble is not when there is an overabundance of realizations, but a dearth. Ramsificationism is at its best when interpreting true theories, terms that actually refer, and so on. But trouble arises when we come to false theories, non-referring terms, and so forth. For here's the thing: It's just as important to correctly interpret such theories, terms, etc., as it is to correctly interpret those that we have yet to find fault with. After all, we no longer believe such theories *because they said* that the world was some way that it was not. Here Ramsificationism faces trouble, but it is more subtle than it may at first appear. The difficulty is *not*, in general, that Ramsificationism understands non-referring terms to be meaningless. After all, although it's true that 'the ether' and 'phlogiston' do not *refer* to anything, it doesn't follow that they are *meaningless*. Quite to the contrary, on the Ramsificationist's picture, the meaning of these terms lies in the descriptions of their functional roles — e.g., 'the ether' roughly means the invisible, continuous plenum through which electromagnetic waves propagate.

[29] That is, suppose a proposed definition τ is indeterminate and, as a result, there are n entities $e_1, \ldots, e_n$ ($n \geq 2$) that can realize τ. In this case, τ is a failure given the ordinary conception of definition. But following the Supervaluationist's lead, the Ramsificationist may adopt a weaker conception of definition, one that only requires that τ's indeterminacy have no effect upon the truth of statements that include it. That is, if the truth value of every sentence Φ containing τ is invariant under assignment of τ to each e_i in the sense that

$$\Phi(\ldots \tau[e_i] \ldots) \leftrightarrow \Phi(\ldots \tau[e_j] \ldots)$$

holds for all $1 \leq i, j \leq n$, then the Ramsificationist may consider this close enough to a definition for her purposes.

So what then is the problem? The problem with this response is that it fails to account for the fact that the descriptions may themselves not be meaningful — at least, meaningful by the Ramsificationist's lights. In particular, this happens when unrealized terms are defined by appealing to descriptions that contain yet further unrealized terms. For without eventually terminating, a chain of unrealized definitions differs little from an arbitrary string of symbols. A second problem is that, at least in some cases, a term may be realized despite possessing a defective description. For example, when Maxwell initially introduced his equations, he explained what they meant in two ways. First, he explained them via Faraday's concept of lines of force. But there are no such things as lines of force. As a contemporary physics text warns (bold face in the text!): "**Electric Field Lines Are Not Real**".[30] And the problem with lines of force isn't simply that, despite being physically meaningful, they are unrealized. Rather, as I noted above, there's a good argument to be made that lines of force were never anything more than a mere heuristic — indeed, Maxwell himself denounced lines of force as 'a purely imaginary substance' despite employing them to explain the meaning of his equations. Maxwell's second way of explaining their meaning was through an analogy with hydrodynamics: he suggests that we think of his equations as describing invisible tubes running through the ether surrounded by ball bearings.

So notice two things here. First, the interpretation of a term may 'factor through' terms that go unrealized or even lack physical significance. Second, the functional role associated with a term may be given only in highly metaphorical language. The Ramsificationist has no story to tell in either case. The latter case is especially problematic. For if the meaning of a term is given in metaphorical language, then the Ramsificationist faces the question: What would an entity have to be like in order to conform with that metaphor? The Ramsificationist not only provides no answer, she *cannot* provide one.

That is to say, although the Ramsificationist could go on to supplement her theory of scientific exegesis with a theory of metaphor, what she cannot do is interpret metaphors using the Ramsification procedure itself. The trouble is that Ramsification procedure can be applied only when a theory's postulate can be taken literally. The Ramsificationist must therefore replace any metaphors with their literal counterparts *before* she can apply the Ramsification procedure. But this means that the Ramsification procedure can never provide a way of interpreting the metaphorical elements of science and, as a result, Ramsificationism cannot fully account for the nature of scientific exegesis.

Limitation 2: The Postulate. A second limitation is that Ramsificationism only provides a way to interpret a theory if we already know the theory's postulate. That is to say, Ramsification only provides a way to interpret a theory if we already know what statements it

[30] Serway and Jewett (2008), 660.

includes. But as we saw above, it often happens that, in point of fact, scientists are unaware of the full range the statements their own theories affirm. Which, in the Ramsificationist's ideology, is just to say that scientists often do not know the postulates of their own theories. So at least from an epistemological point of view, Ramsificationism does little better than Literalism.

But the real problem here is again not epistemological, but metaphysical. Ramsificationism provides no account of what it is to interpret terms, statements, etc., that *occur outside the context of a well-defined theory*. By way of illustration, consider the state of research on electricity and magnetism at the beginning of the nineteenth century. It's not all that much of an exaggeration to say that, during this period, everyone and their mother had a different theory about what electricity and magnetism were, how they behaved, and so on. So the question for Ramsificationism is: During this period what, e.g., does the term 'electric charge' mean? One answer is to say that that nineteenth century physicists didn't actually mean anything in particular. Yet for obvious reasons, this is unattractive. Instead, the Ramsificationist will likely suggest that there are just many different terms—e.g., 'electric charge$_{\text{Ampère}}$', 'electric charge$_{\text{Faraday}}$', 'electric charge$_{\text{Ørsted}}$'—corresponding to each scientist's theory of electricity and magnetism. While this has the advantage that each scientist is saying something meaningful when they invoke the term 'electric charge', this proposal is not without its costs. No two scientists in the period would mean the same thing by 'electric charge', for one.

The Ramsificationist might respond by claiming that this is nothing more than a synchronic version of a familiar problem: the problem of meaning change across theories. (E.g., does 'electric charge' mean the same thing in quantum and classical mechanics?) And the problem of meaning changing across theories is a problem for everyone, not just the Ramsificationist. But the problem here is more pressing. For just what counts as, say, Faraday's theory of electromagnetism? Does it include those only statements that Faraday wrote down in his notebooks? Or does it also include their logical consequences? Does its ideology consist only of those terms Faraday used publicly? Or does it reach further and encompass the whole of Faraday's conceptual scheme? And what of the physical apparatuses found in his lab? Or of his visual experience of watching a compass needle spin? Or his phenomenological state when a current arced to his finger? Are these part of his theory or are they not? Now, don't get me wrong: I'm not expecting anyone to actually be able to go down a list and say precisely what Faraday's theory includes. The point, rather, is that this seemingly pedantic exercise is crucial for Ramsificationism. For given Ramsificationism's top-down picture of exegesis, the meaning of a term like 'electric charge' is fixed by its role in the theory as a whole. As a result, where the boundary of Faraday's theory lies is,

for Ramsificationism, essential both to the epistemological and metaphysical aspects of its interpretation.

Limitation 3: Linguistic Representations. Ramsificationism rests on a limited view of scientific exegesis in yet another way: it only accounts for what it is to interpret representational devices possessed with linguistic structure. While this covers a number of significant types of scientific representations — e.g., scientific theories,[31] models, statements, terms — it fails to account for just as many. Worse, in principal, Ramsificationism cannot account for what it is to interpret many important types of scientific representation.

But first, a few examples of the important representational roles filled by nonlinguistic representations. Such representational devices are ubiquitous in everything from biology to meteorology to atomic physics, especially when dealing with recalcitrant phenomena that elude the systematic treatment of a theory. Examples of such nonlinguistic representations include *photographs* and *diagrams* like Franklin's 'photo 51' of DNA's X-ray diffraction and Crick and Watson's diagram of its double helix structure; *illustrations* such as Galileo's sketches of the moon and Darwin's drawings of finches; *analog models* like wind tunnels and whirlpools; and *computer simulations* like the Monte Carlo simulations used in early investigations of nuclear phenomena.[32] Many other examples could be cited, but the point here is simply that nonlinguistic representations are important vehicles of scientific information and, as such, require interpretation. But Ramsificationism doesn't tell us anything about what it is for, say, a diagram or an equation to represent some part of reality. At best, then, Ramsificationism accounts for the nature of scientific exegesis only in a small handful of cases. It is silent, however, about what is involved in interpreting a great mass of the representational apparatus of science.

The trouble, however, is worse than it first appears. It is not merely that, as a contingent matter of fact, Ramsificationism falls short. The real trouble, rather, is that Ramsification, *in principle*, *cannot* provide a general account of the nature of scientific exegesis. For unless a representational device comes packaged in subject-predicate form, Ramsificationism cannot even begin to say what interpreting it would involve. Therefore, Ramsificationism cannot, in principal, account for what it would be to interpret a sizeable portion of science.

[31] I don't mean to be taking a side here on the question of whether or not scientific theories are linguistic entities—for that, see Hirsch (2017b). For the Ramsificationist, however, theories and their postulates are more or less interchangable. And the postulate of a theory is, in a relatively straightforward sense, a linguistic entity: it is the conjunction of all the sentences the theory affirms.

[32] For excellent discussions of the philosophical issues surrounding the interpretation of computer simulations, see Winsberg (2010) and Galison (1997), chapter 8.

3.3 RAMSIFICATIONISM AND REPRESENTATIONAL ARTIFACTS

Yet even if we set aside all the previous difficulties, there is a more basic problem with Ramsificationism: it just gets the exegetical facts wrong.

Consider electromagnetism once again. Above, we saw one mathematical formulation of Faraday's Law of Induction in terms of vector fields **E** and **B**:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{FLV}$$

There are many other ways of mathematically formulating Faraday's Law (along with the rest of electromagnetism).[33] One way employs what are known as ***potential functions***, φ and **A**. I'll explain what a potential function is shortly, but for the moment let's imagine that we are in the dark and know nothing about potentials other than that, by using them, we may express Faraday's Law as:

$$\nabla \times \left(-\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}\right) = -\frac{\partial}{\partial t}\left(\nabla \times \mathbf{A}\right). \tag{FLP}$$

Observe, that were we to apply the Ramsification procedure to the potential formulation of electromagnetism, then (FLP) would be become a conjunct in its postulate:

$$(\mathfrak{P}_{EM}) \qquad \ldots \text{ and } \left(\nabla \times \left(-\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}\right) = -\frac{\partial}{\partial t}\left(\nabla \times \mathbf{A}\right)\right) \text{ and } \ldots$$

Thus, when we existentially generalize over φ and **A**, we come to have a Ramsey sentence of the form:

$$\ldots \exists x_1 \exists x_2 \ldots \left[\ldots \text{ and } \left(\nabla \times \left(-\nabla x_1 - \frac{\partial x_2}{\partial t}\right) = -\frac{\partial}{\partial t}\left(\nabla \times x_2\right)\right) \text{ and } \ldots\right].$$

Now here's the key observation: If we took this Ramsey sentence as our interpretation of electromagnetism then we would be taking electromagnetism to posit the existence of electric and magnetic potentials. But this is a fundamental misunderstanding. For the thing about potentials is that they are representational artifacts. As one electromagnetism text explains:

> [The electric] potential as such carries *no real physical significance*, for at any given point we can adjust its value at will…In this sense it is rather like altitude: If I ask you how high Denver is, you will probably tell me its height above sea level, because that is a convenient and traditional reference point. But we could as well agree to measure altitude above Washington D.C., or Greenwich, or wherever. That would add (or, rather, subtract) a fixed amount for all our sea-level readings, but it *wouldn't change anything about the real world.* The only quantity of intrinsic interest is the difference in altitude between two points, and that is the same whatever your reference level.[34]

[33] See Warnick and Russer (2014) and Hirsch (2015), section 3.

[34] Griffiths (1999), 80. My emphasis.

And similarly with the magnetic potential.[35] So it would be a mistake to understand (FLP) as positing an electric and a magnetic potential. Indeed, doing so would lead us to conclude that (FLP) and (FLV) are representationally *in*equivalent. But any physicist will tell you that (FLP) and (FLV) are just different ways of saying the same thing. So when representational artifacts are involved, Ramsificationism just gets the exegetical facts wrong.

How might the Ramsificationist respond? One way is to bite the bullet and maintain that the physicists just are confused about the content of electromagnetism. But this is hardly an attractive option. For one thing, it smacks of precisely the armchair imperialism that the Ramsificationist's deference to science is meant to oppose. For another thing, as I shall argue below, to ignore the judgments of actual scientists is to overlook an essential aspect of scientific exegesis. So if the Ramsificationist is to respond, it must be by resisting at the point when we existentially generalized over φ and $\mathbf{A}$. But this flies in the face of the Ramsificationist's whole proposal. For if Ramsificationist fails to generalize over φ and $\mathbf{A}$, then she has no story about why (FLP) is not merely an equation of pure mathematics but instead a full-blooded, physically significant representation of reality.

4 EXEGETICAL ANGST OR INTERPRETIVE INSIGHT?

It's not hard to see that the Problem of Representational Artifacts presents an *epistemological* obstacle to understanding what science says. Look at it this way: On the one hand, we might mistake a representational artifact for a substantive aspect of science. And were we to do so, we would understand science to be saying something that it is not. On the other hand, we might err in the opposite direction by mistakenly understanding some substantive aspect of the language to be a mere representational artifact. In this case, we would fail to recognize something that science is, in fact, saying about the world. So The Problem of Representational Artifacts raises the possibility of two equally egregious exegetical errors: mistakenly ascribing physical significance to a representational artifact, and erroneously dismissing physically significant features as mere representational artifacts. In either case, the effect is the same: failure to understand what science says about the world.

While there's no denying the importance of this epistemological obstacle, I want to argue that The Problem of Representational Artifacts' true importance lies in what it reveals about the nature of scientific exegesis. And what it reveals, I'm going to argue, is that scientific exegesis factors into a number of distinct tasks, several of which the Literalist and Ramsificationist simply ignore.

[35] It's a little more mathematically complicated in this case because the magnetic potential is a vector- valued function and not scalar-valued like the electric potential. The important point, however, still applies: the magnetic potential has 'no real physical significance'.

4.1 A CARTOGRAPHIC ANALOGY

One problem with discussing The Problem of Representational Artifacts is that, owing to the abstractness of actual scientific examples, it can be hard to put one's thumb on The Problem's essential features. But The Problem of Representational Artifacts is not unique to scientific exegesis. Quite to the contrary, any method of communication possesses attendant aspects that are incidental to the information it conveys. The language in which a message is expressed, for instance, or the typeface and color of ink with which it is printed, are paradigms of representational artifacts. So let's take a brief detour and consider an everyday example where The Problem of Representational Artifacts arises.

Imagine two hapless explorers, Alice and Bob, lost in the woods. While they had the foresight to print out maps, neither remembered to bring the maps' keys. What do Alice and Bob need to understand in order that they might use the maps to get out of the woods? In the first place, Alice and Bob must decide which of the maps' various markings, patterns, and so forth, are representationally significant. Do the squiggles on Alice's map signify anything? Are the grid lines on Bob's map important? And just just as importantly, they also need to decide which features of their maps are representationally insignificant. After all, each map possesses mere stylistic flourishes, printing blemishes, and any number of other markings that do not signify anything about the terrain. We know that those on Alice's map are topographical contour lines and that Bob's grid is merely an artistic decision by the cartographer, but if Alice and Bob didn't realize this they could be in serious trouble. If, for instance, they mistook the squiggles for decorative elements they might find themselves a top a cliff with no way forward. On the other hand, if they took the grid lines to signify roads, they might walk for hours in the wrong direction looking for a highway. So the first challenge Alice and Bob face is:

Cartographic Challenge 1: Determine *which* of the maps' features are significant and *which* are mere representational artifacts.

But even if Alice and Bob figure out which features of their maps are significant, they face further challenges. Their maps do them no good unless they can also determine what those features purport to signify about the terrain. Are the squiggles intended to represent roads, or perhaps streams, or—as we know they do—elevation? Knowing the answer is, again, crucial to correctly interpreting the maps. So the second challenge they face is:

Cartographic Challenge 2: Determine *what* the maps' significant features purport to signify.

Yet even if they've successfully navigated these first two challenges and determined that the squiggles represent elevation they still face the question of *how* these squiggles are used to

encode elevation facts. Is it that the squiggles denote specific elevations (e.g., one of squiggle denotes 100 meters, another 500, etc.)? Or is it that the squiggles represent changes in elevation (e.g., the space between two squiggles represents a change of, say, 100 meters)? Thus, the third challenge they face is:

Cartographic Challenge 3: Determine *how* the maps' significant features represent the terrain.

While all three cartographic challenges are exegetical in nature—they concern *what the maps say about* the terrain—each involves a fundamentally different task. The second challenge closely resembles the sense of exegesis familiar to Literalist and Ramsificationist: it amounts roughly to translation at the lexical level. It involves translating cartographic terms into topographic ones—e.g., 'contour line' means 'elevation'. The first and third challenges, however, are quite different. They do not involve translating the lexicon of one language into another, but translation of a more primitive sort: translation between two fundamentally different types of representation, from a map that lacks subject-predicate structure into a collection of linguistic devices with such structure. This more primitive sort of translation is one manifestation of what, to transplant a Quinean term, I shall call ***regimentation***.[36]

Regimentation is the process of associating a representational device and its features with elements of one's physical language. Consider how this goes in the case of maps. The first step is to determine *which* of the map's features are significant and which are mere artifacts. Alice might point at the contour lines and suggest, "Lines like these ones are important." Bob may further suggest, "The spacing between the lines is also significant." Yet even after determining *which* features of the map are significant, Alice and Bob still must say *how* these features encode information and *what* information they encode. The main challenge in saying *how* these features encode information is to establish a correspondence between these features and the syntactic categories of Alice and Bob's physical language. For instance, do the maps' contour lines correspond to properties, or to relations, or to something else entirely? And further, how do the contour lines correspond to this syntactic category? Alice may thus suggest, "The contour lines encode relational facts." And Bob may add, "In particular, it is the density of the lines that encodes such facts." Only once Alice and Bob have successfully regimented their maps can they then turn to the familiar lexical stage of exegesis: translating claims about the representationally significant aspects of their maps into statements about the physical features of the terrain.

Now, it's important to observe that regimentation is *logically prior to* the lexical translation familiar to the Literalist and Ramsificationist. Before Alice and Bob even reach the

[36] Cf., Quine (1960), chapter 5.

stage of lexical translation, they must successfully regiment the maps: they must isolate the representationally significant features while setting aside the mere stylistic flourishes, printing blemishes, and the like. If, however, they fail at this earlier stage, then their exegetical efforts at the later lexical stage will be foiled before they've even begun. To assign *any* significance to the blemish is to take the maps to say something that they do not. So, more generally, exegetical errors during regimentation trickle down and derail the later stages of scientific exegesis.

Further, observe that regimentation gives rise to a great deal of *exegetical freedom*—or perhaps better, a great deal of *opportunity for exegetical failure*. For the thing about regimentation is that it can involve associating what might be called 'undifferentiated' representational devices with 'differentiated' ones. This distinction is tricky to spell out, but it is not unintuitive. A sentence like "There are electrons" is differentiated in a relatively straightforward sense: it contains exactly three words. Maps, however, are undifferentiated: they do not come 'prepackaged' with a single, well-defined collection of cartographic features. We shall revisit this distinction shortly, but for the moment I'm going to assume that we have a basic idea of what the distinction is getting after. Now, what we want to observe is that undifferentiated devices introduce an extra kink into the exegetical process. Because their maps are undifferentiated, it is not as though Alice and Bob can regiment them by simply running down a finite list of cartographic features, circling the representationally significant features and then crossing off the remainder as artifacts. Rather, the first step in regimenting the maps involves 'carving' them up into such a list of features. But here's the thing: there is no limit to the possible patterns of pigment that Alice and Bob might pick out as representationally significant. So, owing to their lack of differentiation, the maps may be 'carved up' in countless ways. Of course, this is not to say that all such carvings are on an equal exegetical footing. To the contrary, there is at most one correct carving: namely, whichever carving singles out all and only the representationally significant cartographic features for further interpretation. The need to carve up a representational device means, more generally, that there is no limit to how it might be interpreted. But the overwhelming majority of these interpretations go wrong at the very first step.

4.2 REGIMENTATION AND THE TASK OF SCIENTIFIC EXEGESIS

Bringing the discussion back to science, my contention is that scientific exegesis differs little from interpreting maps when it comes to the need for regimentation. In a moment I'll argue why this is an apt analogy, but first allow me to highlight its key points. The main thing is that regardless of whether R is a scientific representation or a map, there are three phases to interpreting R:

Phase 1. Determining *which* of *R*'s features are significant and which are mere representational artifacts.

Phase 2. Determining *what R*'s significant features purport to signify.

Phase 3. Determining *how R*'s significant features encode information about the phenomena they purport to signify.

Moreover, just as there is a logical priority amongst these three phases of interpretation in the case of maps, so too is there in the case of scientific representations: if, for instance, we incorrectly identify which of *R*'s features are substantive and which are representational artifacts, then our exegetical efforts in the latter two phases will be thwarted before we've even begun.

So just how strong is the analogy between scientific exegesis and interpreting maps? Let's begin by considering the analogy in the case of non-linguistic representations and then move to the harder case of linguistic representations.

The Case of Non-Linguistic Representations

When it comes to interpreting non-linguistic representational devices, especially visual devices like diagrams, photographs, and illustrations, there is little difference to interpreting a map. The key thing is that these devices do not come tidily packaged in subject-predicate form any more than maps do. Photographs are composed of paper and ink, or strings of zeros and ones, not of noun- and verb-phrases. But if photographs represent objects and properties—and surely many do—then interpreting photographs begins with identifying which patterns of pigment mark out objects, which packs of pixels signify relations, and so forth. Or to impose an imaging metaphor, interpreting a photograph requires separating signal from noise.

With other types of non-linguistic representations, while the analogy to interpreting maps may be less overt, it is certainly no less apt. Interpreting computer simulations, for instance, involves all three exegetical phases despite the fact that computer simulations are not pictorial representations like maps or photographs. And while there may be no printing blemishes or image noise to worry about when interpreting computer simulations, this hardly means that there are no representational artifacts. Quite to the contrary, simulations contain many representational artifacts, though of a different sort: they are the aspects of a simulation that merely reflect the assumptions of one's algorithm rather than signifying anything significant about the physical phenomena being simulated.[37]

More generally, interpreting a non-linguistic representation is analogous to interpreting a map in the sense that it involves more than translation at the lexical level: it also involves

[37] For an excellent discussion of just this point, see Winsberg (2010).

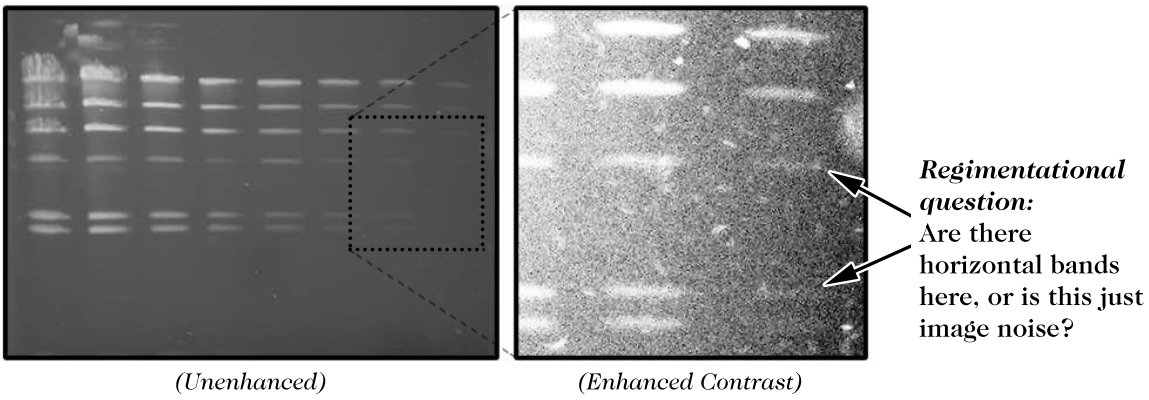


Figure 3.1 *Regimentation in gel electrophoresis.*

determining *which* features are representationally significant and *how* these features encode information about the phenomena they purport to represent. The case of interpreting a photograph will be help to bring out some of the details of this analogy. Consider the photograph in figure 3.1. This photograph was the product of an experiment that I ran to test the outcome of modifying E. coli DNA using a technique known as 'gel electrophoresis'. The technique involves placing the DNA into one end of a porous gel (imagine seaweed-flavored Jello) and then applying a small electric charge to the other end. Now, because DNA and other proteins are negatively charged, they are attracted to the opposite side of the gel. But the DNA segments do not move through the gel at the same rate. Rather, shorter segments move more quickly than longer ones owing to the fact that they can navigate through the gel's pores more easily. After 'running the gel' for some time (around an hour in this case), the electric charge is turned off and the gel is then dipped in ethidium bromide to stain the DNA segments so they will glow when placed under a florescent light. The photograph in figure 3.1 is of the end result of this process.

Given all this, the exegetical question is: What does this photograph indicate about E. coli DNA? Now, the answer obviously depends upon physical details of the experimental setup—e.g., how porous the gel was, how long it was run for, what sort of catalyst was used to 'chop up' the E. coli DNA, and so on. But supposing we know all this, three further facts are crucial to correctly interpreting the photograph. First, the horizontal bands indicate segments of E. coli DNA. Second, because shorter segments of DNA can more easily navigate the porous gel, they travel further than longer segments (indicated in this case by being closer to the bottom of the image). Third, the bands' brightness varies in proportion to the amount of DNA present (e.g., going from left to the right, each column contains smaller amounts of DNA). In other words, all three exegetical phases are required to correctly interpret this photograph:

Phase 1. *Question: Which* of the photograph's features are significant and which are mere representational artifacts? *Answer:* The horizontal bands, their vertical position, and their brightness.

Phase 2. *Question: What* do these features purport to signify? *Answer:* DNA segments, the length of those segments, and the quantity of those segments.

Phase 3. *Question: How* do these features encode information about the phenomena they purport to signify? *Answer:* A band's vertical position indicates the length of a DNA segment, with shorter segments being closer to the bottom of the image; the brightness of a band indicates the quantity of DNA, with brighter bands indicating more DNA.

While this knowledge affords us with a *general* recipe for interpreting photographs of electrophoresis gel experiments, one crucial exegetical question remains to interpret this *particular* photograph: How many horizontal bands does the photograph contain? Here, The Problem of Representational Artifacts looms large. For the outstanding question is a question what to count as a horizontal band and what to write off as mere image noise. While some cases like those on the far left are fairly clearcut, those on the right are not. In particular, when we turn to the bottom-right of the photograph, it's far from obvious what, if anything, to count as horizontal band. Worse, as I adjusted the image contrast, I discovered that different methods of adjustment appeared to give different answers (see figure 3.2): with method A (no adjustment) there appears to be no horizontal bands in the indicated region; with methods B and C, one band; and with method D, two bands. And this is troubling, for what the original photograph indicates about E. coli DNA depends upon how many horizontal bands are in this region.

The Case of Non-Linguistic Representations

While it's not hard to feel the analogy's force in cases of non-linguistic representations like photographs and simulations, it appears somewhat more tenuous when we turn to linguistic representations.

When we are presented with a representational device like the sentence "There are electrons", it's tempting to think that there is no need for further regimentation. And were this so, then scientific exegesis would require nothing more than the second, lexical phase—e.g., nothing more than defining the term 'electron'. So one of the main complaints against the map analogy is that it only applies in cases non-linguistic representations. For the reason that such representations require regimentation, the complaint goes, arises from the fact that they are undifferentiated and so require that we first carve out some set of representationally significant features. Linguistic representations, however, come to us carved

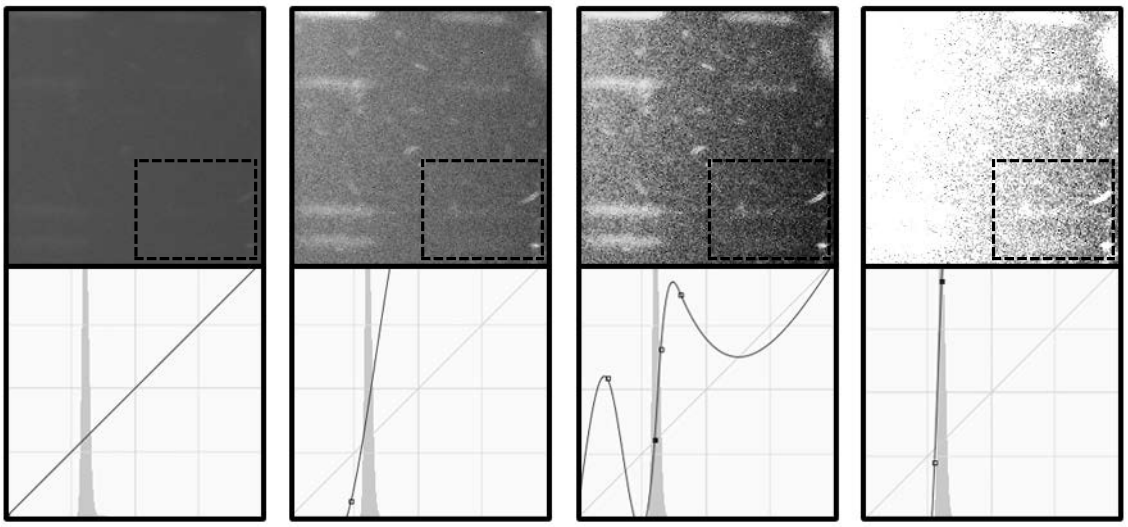


Figure 3.2 *Above: Results of different contrast adjustments. Below: Histograms corresponding to the contrast adjustment.*

up into a collection of exegetically salient elements; they come regimented. So, the complaint concludes, unlike non-linguistic representations, linguistic representations require no regimentation.

Here is where The Problem of Representational Artifacts begins to reveal its deeper significance, as not merely a cause for exegetical angst, but also as a source of interpretive insight. For what The Problem of Representational Artifacts shows, I will now argue, is that regimentation is required even in cases of linguistic representations, even when a representational device appears on its face to have already been regimented. By way of illustration, consider the potential formulation of electromagnetism. If we so desired, we could write out the potential formulation of Faraday's Law of Induction,

$$\nabla \times \left(-\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t} \right) = -\frac{\partial}{\partial t} \left(\nabla \times \mathbf{A} \right),$$

in natural language, giving us a cumbersome monstrosity along the lines of:

(FLP) There are electric and magnetic potentials φ and $\mathbf{A}$ such that the rotation of φ's dispersion minus $\mathbf{A}$'s change over time is equal to the amount that $\mathbf{A}$'s rotation is changing over time.

The crucial observation is that, despite possessing a subject-predicate structure, (FLP) still requires regimentation. For as we saw above, φ and $\mathbf{A}$ are representational artifacts. Correctly interpreting (FLP) therefore requires that we recognize that the potentials do not denote physical entities, properties, etc. Further, correctly interpreting (FLP) also requires us to recognize that it encodes information about entities that are not even explicitly men-

tioned: the electric and magnetic fields. And further still, a correct interpretation also requires recognition of *how* φ and **A** are used to encode this information: roughly, the value of potential at a location ℓ encodes *the difference in* charge distribution between ℓ and some arbitrarily chosen point of reference. As a result, neither the potentials themselves nor their specific values are representationally significant; the only representationally significant quantity is *the difference in* a potential's value. So despite having a subject-predicate structure, correctly interpreting (FLP) still involves all three exegetical phases: it requires recognizing (i) that φ and **A** are not directly significant; (ii) that differences in φ and **A** are physically significant; and (iii) these differences signify how the strengths of the electric and magnetic fields vary between locations.

So The Problem of Representational Artifacts foists upon us a need for regimentation even in cases of linguistic representations like (FLP) that are, in some sense, already differentiated. It's important to see that this need for regimentation arises precisely because there is a need to identify which features of the representations are significant and which are mere artifacts. And this suggests that linguistic and non-linguistic representations differ little in point of requiring regimentation. Indeed, despite the fact that (FLP) appears prepackaged into thirty-some exegetically salient elements, it can be carved up in many different ways just like the photograph in figure 3.1. Observe how this happens. For one thing, nothing about φ and **A** requires that they be representational artifacts rather than directly significant aspects of (FLP). So there are ways of carving up (FLP) that take φ and **A** to be representational artifacts and others that do not. More importantly, it is the question of *how* φ and **A** encode information that leads to a proliferation of possible carvings. Take φ as an example. Mathematically speaking, it is what is known as a scalar field—that is, a function that assigns real numbers to every point in some space. In regimenting φ, however, we do not associate it with some other *scalar* field, but rather a *vector* field, **E**. And it is this vector field that we take to correspond in a direct sense to the character of the electric field. So note four things. First, in regimenting φ, we are associating with it some intermediary mathematical entity, **E**. Second, this intermediary mathematical entity is *never explicitly mentioned* in (FLP), nor anywhere else in the potential formulation of Maxwell's equations. Third, φ and **E** do *not* have the same mathematical structure—one is a scalar field, the other a vector field. Fourth, there is no general requirement that scalar fields encode information through differences in their values; quite to the contrary, the numerical values of some scalar fields (e.g., like those that represent continuous mass distributions) are, in fact, physically significant.[38] To summarize: In regimenting φ we are associating it with a mathematical entity **E** that (i) is not explicitly mentioned in (FLP) nor anywhere else in the potential formulation

[38] That is, ignoring the question of Absolutism versus Comparativism about physical quantities. Cf., Dasgupta (2013).

of Maxwell's equations, (ii) that has a different mathematical structure than φ, and (iii) in general has no necessary connection to φ. There are, in other words, few if any constraints on how we may carve up φ.

The moral here, more generally, is that linguistic and non-linguistic representations differ little from an exegetical point of view: not only do both types of representation require regimentation, both can be 'carved up' in many different ways. In light of all this, I submit that the map analogy is every bit as apt in the case of linguistic representations as it is in the case of non-linguistic ones.

4.3 REGIMENTATION AND THE NATURE OF SCIENTIFIC EXEGESIS

One might protest that even if both linguistic and non-linguistic representations require regimentation, the two modes of representation differ significantly in *how they represent*. For instance, non-linguistic representations are sometimes said to 'represent via isomorphism or structural similarity' rather than by 'represent via naming' as linguistic representations do.[39] And if this is so, it might suggested that the analogy with maps is not nearly as telling in the case of linguistic representations as I have suggested. This is an important suggestion and well worth considering. However, it would take us too far afield to fully explore now, so I would like to set this complaint aside until the final chapter.[40] Still, I should say a few words about this complaint as it points to an important connection between regimentation and the nature of scientific exegesis.

So far, I've primarily focused on The Problem of Representational Artifacts and the resulting need for regimentation from an *epistemological* point of view, as a problem for *determining* what science says and a task required to overcome it. In particular, we've seen that both Literalism and Ramsificationism lead us to misinterpret science in instances of representational artifacts. This problem can be seen as an epistemological fault: neither Literalism nor Ramsificationism provide an adequate account of how to determine what science says. Yet while this epistemological problem is important, The Problem of Representational Artifacts' true significance lies in what it reveals about *the nature of* scientific exegesis. And what it points to is the fact that Literalism and Ramsificationism both possess an overly simplistic conception of the nature of scientific exegesis. Recall that one

[39] Cf., Lewis (1986), chapter 3.

[40] As a preview, my contention is going to be that this complaint looks at things the wrong way around. Because both linguistic and non-linguistic representations require regimentation, there is no significant distinction in how they represent. To say that a linguistic representation L 'represents via naming' is to say that, if the world were as L says it is, then there must be a correspondence between the significant features of L and elements of the world. But as we have just seen, this is no different in the case of non-linguistic representations. In regimenting a representational device R, whether R is a linguistic or non-linguistic representation, we are isolating the features of R that must correspond to elements of the world if the world is as R says it is. In this sense, non-linguistic representations 'represent via naming' every bit as much as linguistic representations.

necessary condition for interpretive adequacy according to both Literalism and Ramsificationism was that an interpretation of a statement (theory, etc.) entailed the statement's purely structural commitment. (In this sense, Ramsificationism just is a form of Literalism, albeit one with a relaxed standard of definition.) The crucial thing about this necessary condition is that it assumes that there is a *tidy connection between ideology and ontology*—in its most naïve form, predicate symbols correspond to properties and relations, constants to individuals, and so forth.

Having come to appreciate that scientific representations contain physically insignificant artifacts and so require regimentation, we can now see that there is no straightforward connection between scientific ideology and ontology. Consider again the example of (FLP). The first-order exegetical question is: What would the world have to be like to be as (FLP) says it is? According to both the Literalist and Ramsificationist, any adequate answer must minimally acknowledge that the world would have to contain physical entities denoted by the potential symbols 'φ' and '**A**'. But as we have seen, these symbols are representational artifacts. They not only have no denotation, they do not *even purport to* denote physical entities. In this sense, potentials are like compass roses on maps—while they are used to encode information about the physical world, they themselves are not intended to directly correspond to any part of the world. The result, as we have seen, is that the correct interpretation of (FLP) involves regimentation. And regimenting (FLP) involves associating the potentials with entities that are neither explicitly mentioned in (FLP), nor have the same mathematical structure as the potentials themselves, and so on. In short, the connection between (FLP) and the world is far more complicated than anything recognized by either the Literalist or Ramsificationist.

So as I see it, the true significance of The Problem of Representational Artifacts lies not in the fact that it presents an epistemological problem, a problem for determining what science says. Rather, its importance lies in the fact that it reveals the connection between scientific representations and the world to be far more complicated than is commonly appreciated: *The Problem of Representational Artifacts shows that there is no tidy correspondence between the ideology and ontology of science.*

4

LANGUAGE, STRUCTURE, AND SCIENTIFIC EXEGESIS

Questions about the semantics of science were all the rage during the heyday of logical positivism. Philosophers asked if theoretical terms like 'gene' and 'electron' were meaningful.[1] They puzzled over how ideology borrowed from pure mathematics like '10-dimensional' might come to have physical significance.[2] They wondered aloud if seemingly different theories like special relativity and Lorentz's ether theory were really just saying the same thing.[3] And so on.

In the decades following positivism's eclipse, however, many philosophers came to downplay the importance of such semantic inquiry. While the reasons for this are undoubtedly complex, the long and short of it is this: Over time, the standard textbook problems attributed to positivism—e.g., the trouble formulating the verification criterion of meaning,[4] the untenability of the observable-theoretical dichotomy,[5] the vagueness of the analytic-synthetic distinction,[6] to name just a few—came to be seen as symptoms of a deeper disease: a fetish for language. That is to say, many philosophers came to believe that the positivism's problems were, at bottom, products of an unduly linguistic conception of science.

This complaint, what I shall call ***The Linguistic Objection***, is ubiquitous. For instance, Bas van Fraassen writes:

> My own view is that empiricism is correct, but could not live in *linguistic form the positivists gave it.* … The logical positivists, and their heirs, went much too far in this attempt to turn philosophical problems into *problems about language.* In some cases *their linguistic orientation had disastrous effects in philosophy of science.*[7]

[1] See Carnap (1936).
[2] See Reichenbach (1965), chapter 4.
[3] See Schlick et al. (1979), 153–189.
[4] See Hempel (1950).
[5] See Maxwell (1962).
[6] See Quine (1951).
[7] van Fraassen (1980), 3-4. My emphasis.

Yet for all the protest to positivism's linguistic orientation, its alleged language-dependent conception of science, etc., surprisingly little has been said about why these notions are problematic, and even less about what they amount to. Just what does it mean to say, e.g., that the positivists' possessed a linguistic conception of science?

A closer look at the The Linguistic Objection is therefore overdue. And not merely because it is a prominent objection, but more importantly, it has shaped many core assumptions underlying the contemporary philosophy of science. For here's the thing: Having pinned positivism's problems on its linguistic orientation, many philosophers came to feel that the remedy was to turn away from language. Take as an example the case of scientific theories. The prevailing wisdom is that the positivists' conception of theories went wrong by viewing theories as linguistic entities, as roughly sets of sentences. The trouble, we are told, is that this confuses the *language used to express scientific theories* with the *theories themselves*. Frederick Suppe, for example, writes:

> Such an approach [the positivist conception of theories] can provide a detailed analysis of the characteristic features of theory formulations, but unless one assume that analogues to these features are distinctive features of theories themselves; *such an analysis reveals nothing about what is characteristic of theories except that their formulations have certain characteristics*.[8]

And if this is so, if positivism's problems were products of having a language-dependent conception of theories, then aren't language-*in*dependent conceptions the only viable alternatives?

Many philosophers have thought so. For instance, Suppe writes (continuing the previous passage), "These considerations suggest that a *direct examination of theories themselves* is likely to result in a more detailed and accurate analysis of the structure of theories than is possible if one concentrates one's attention on theory formulations."[9] Encouraged by such suggestions, much of the philosophy of science came to rest on the assumption that scientific theories are language-independent entities that could be examined directly. And this, in turn, led to the widespread perception that language-independent conceptions of science avoid the problems in which positivism had become mired. Thus, Suppe claims elsewhere that:

> The [postpositivist conception of science] has been very successful. It is widely accepted with remarkably little published criticism of it—none fundamental or fatal. It also avoids the problems that scuttled the [positivist conception]: Correspondence rules [roughly, interpretive rules] are avoided. Observational/non-observational dis-

[8] Suppe (1974), 221–222.
[9] Suppe (1974), 221–222.

> tinctions are not required. Formalism artifacts that plagued positivism have not occurred. It individuates theories better than positivism did.[10]

And if this is so, then it's a short step to seeing the semantics of science as little more than a sideshow. After all, given that these issues ostensibly arise from focusing on language, if such a focus is misguided then seemingly the issues are too. Or as van Fraassen writes:

> Perhaps the worst consequence of [the positivist conception of theories] was the way it focused attention on *philosophically irrelevant technical questions*. It is hard not to conclude that those discussions of axiomatizability in restricted vocabularies, 'theoretical terms', Craig's theorem, 'reduction sentences', 'empirical languages', Ramsey and Carnap sentences, were one and all off the mark—*solutions to purely self-generated problems, and philosophically irrelevant. The main lesson of twentieth-century philosophy of science may well be this: no concept which is essentially language-dependent has any philosophical importance at all.*[11]

While I do not pretend that positivism was free of difficulties, it seems to me that this linguistic diagnosis misses the mark.

I'm going to argue that contrary to popular philosophical opinion, the semantic issues that occupied the positivists have little to do with language—at least, language in anything like the ordinary sense of the term. Rather, they concern the relation of theory and world; that is to say, the nature of representation. My contention will therefore be that the key place where postpositivist conceptions of scientific theories depart from that of the positivists is not, as commonly believed, that the latter is language-dependent while the former are not. The fundamental difference is, instead, a difference in how the positivists and postpositivists understand the nature of representation. This is not merely a point of historical interest. Quite to the contrary, it suggests that postpositivists' dismissal of the alleged linguistic issues that occupied the positivists was hasty. For if these issues do in fact concern the nature of representation, then they are issues that arise on any conception of science, regardless of how linguistically oriented it may be. After all, even if theories are not linguistic objects, they are representational ones. And this means that we can (and should) ask what theories represent and how they represent it.

1 INDIVIDUATION AND ABSTRACTION

Overwhelmingly, complaints about positivism's linguistic orientation have been directed specifically at its conception of ***scientific theories***, sometimes known as 'the received

[10] Suppe (2000), 105.
[11] van Fraassen (1980), 56. My emphasis.

view of theories' or 'the syntactic view of theories'.[12] Condensed to lapidary slogan, it is the view that *scientific theories are sets of sentences.* While many of the details are irrelevant for present purposes, there are few things we should note.

First, the positivists borrowed their conception from mathematical logic where a theory is (roughly) any set of sentences closed under implication.[13] The positivists adapted this mathematical concept into a full-blown conception of scientific theories by imposing additional requirements on the sentences of a scientific theory. To comprise a scientific theory, a set of sentences had (again, roughly) to factor into three parts: (i) a set of sentences expressed in an ***observation language*** whose vocabulary have empirical consequences, e.g., 'red', 'heavier than', and so on; (ii) a set of sentences expressed in a ***theoretical language***, e.g., whose vocabulary includes terms like 'electron', 'gene', so forth; and (iii) a set of sentences called ***correspondence rules*** linking the vocabulary, statements, etc., of the theoretical language to the observation language.

This may seem arcane, but it is intended to explicate an intuitive picture. On the one hand, there is the empirical or physical phenomena investigated by science. On the other hand, there are the abstract systems of signs and symbols employed by scientists to represent and theorize these phenomena. But for these abstract systems to be of any use, there also needs to be rules for translating talk about them into talk of the phenomena. So correspondence rules are, in a manner of speaking, the conduits through which theoretical elements imbibe empirical significance.

Thus, on the positivist conception theoretical vocabularies and languages are, *sans* correspondence rules, nothing but arbitrary collections of symbols and formal calculi. As Carnap explains:

> [W]e can first construct a calculus and *then lay down the interpretation intended in the form of semantical rules, yielding a physical theory* as an interpreted system with factual content…Thus, for instance, a calculus of mechanics of mass points can be constructed. Some predicates…are taken as specific primitive signs, and the fundamental laws of mechanics as axioms. Then semantical rules [i.e., correspondence rules] are laid down stating that the primitive signs designate, say, the class of material particles…*By the interpretation, the theorems of the calculus of mechanics become physical laws*…The relation of this theory to the calculus of mechanics is entirely analogous to the relation of physical

[12] Despite my inveterate use of phrases of the form "*the* positivist conception of…" I do not mean to suggest that there is a single positivist conception of science, or of anything else for that matter. Logical positivism, like any other broad philosophical program, comes in a variety of flavors. But insofar as we are concerned with a philosophical question—what is the linguistic objection to positivism and how has it impacted postpositivist conceptions of science—it will be helpful to set aside the various and sundry points of internecine dispute and focus instead on the broad positivist picture, one roughly along the lines of that presented in part 5 of Carnap (1995) and what Suppe (1974) calls 'the final version of the Received View' (50).

[13] That is to say, for a language $\mathcal{L}$, a set T of $\mathcal{L}$-sentences is a theory iff T contains every $\mathcal{L}$-sentence entailed by the members of T.

> to mathematical geometry. The customary division into theoretical and experimental physics corresponds roughly to *the distinction between calculus and interpreted system*.[14]

Take for example an equation like "$F = ma$". As part of the theoretical language of mechanics, it is nothing more than a formal expression of pure mathematics on a par with arbitrary differential equations like "$G = nb$" and "$X = yz$". So what sets "$F = ma$" apart? For the positivists, correspondence rules furnish the answer. As Carnap put it elsewhere, it is through correspondence rules that "the blood of empirical reality flows … into the veins of the hitherto empty schemata, thereby transforming them into a full-blown theory."[15]

1.1 INDIVIDUATING THEORIES

Of the complaints to the positivist conception's linguistic orientation, ***The Individuation Objection*** is both the most prominent and the most straightforward. The basic complaint is that linguistic entities, and sets of sentences in particular, are *too fine-grained* to serve as scientific theories. Thus, Frederick Suppe argues: "theories are not linguistic entities and thus theories are *individuated incorrectly*."[16] The trouble is that if theories are sets of sentences, then it would be impossible to express a theory in more than one way. As Giere writes:

> Since it seems that theories are things one finds written down in textbooks, among other places, one is inclined to look for something "linguistic" to be the theory itself. But one cannot identify the theory of mechanics with any particular set of sentences. *That would make the translation of a textbook from English into French a different theory, which seems absurd*.[17]

Now, I think it's fair to say that this is a decisive objection to the understanding scientific theories as sets of sentences. For even setting aside the possibility of formulating theories in *different* languages, surely everyone must admit that any language is rife with inconsequential variation. The theory of mechanics may be expressed in the active voice, but the passive voice may also be used to express it. And yet, by any reasonable standard, there is no substantive difference between the two; they are one and the same theory. However we conceive of scientific theories, some degree of variation must be allowed. So identifying theories with set of sentences is untenable.

Despite this, I do not think that The Individuation Objection gets at the heart of what philosophers have found objectionable in the positivist conception of theories. Quite to the contrary, as objections to the positivist conception go, The Individuation Objection is

[14] Carnap (1939), 198–199. My emphasis.
[15] Carnap (1927), 372–373. Translation from Ben-Menahem (2006), 185.
[16] Suppe (2000), 103. My emphasis.
[17] Giere (1988), 84. My emphasis.

rather shallow. For one thing, there is a great deal of evidence to suggest that, as an objection to the positivist conception, The Individuation Objection is a strawman: the positivists weren't so naïve, or so in the grips of some philosophically preconceived idea of what scientific theories should be, to maintain that French and English formulations of mechanics are different theories. Quite to the contrary, time and again we find Carnap and others insisting that there multiple ways of saying the same thing.[18]

But for the sake of argument, let's suppose that the 'official' positivist view maintained that French-mechanics really was a distinct theory from English-mechanics. In this case, The Individuation Objection would weigh decisively against the positivist conception of theories. Yet—and here's where The Individuation Objection's shallowness shows—it would only be an objection to the letter of the positivist conception, not its spirit. Observe that The Individuation Objection tacitly targets a notion that we might call 'strict language-dependence':

> A conception of scientific theories is ***strictly language-dependent*** if, for any theories T and T', any linguistic difference between T and T' implies that T and T' are different theories.

In essence, The Individuation Objection aims to show that strictly language-dependent conceptions of theories draw distinctions where really there are none.

The trouble with the objection, however, is that just because a conception of theories is strictly language-dependent doesn't mean that it draws distinctions at the drop of a hat and hence succumbs to The Individuation Objection. Quite to the contrary, with only minor modifications, any strictly language-dependent conception of theories can easily avoid The Individuation Objection. Let $\mathcal{C}$ be a strictly language-dependent conception of theories, and let T and T' be $\mathcal{C}$-theories that differ only inconsequentially (e.g., T is the active voice while T' is in the passive, or T is in English and T' in French, etc.). But in this case, there is a nearby, more coarse-grained conception of theories, $\mathcal{C}^\star$, where a $\mathcal{C}^\star$-theory is an equivalence class of $\mathcal{C}$-theories closed under differences of inconsequential variation. That is, for any $\mathcal{C}$-theory S, there is a $\mathcal{C}^\star$-theory $[\,S\,]$ defined as

$$[\,S\,] =_{df} \left\{\, S' \in \mathcal{C}\text{-theories} \;\middle|\; S' \text{ differs inconsequentially from } S \,\right\}.$$

For instance, on the $\mathcal{C}^\star$-conception, the English and French formulations of mechanics are one and the same theory. Thus, although T and T' are distinct $\mathcal{C}$-theories, because they differ only inconsequentially, $[\,T\,]$ and $[\,T'\,]$ are one and the same $\mathcal{C}^\star$-theory. Observe that $\mathcal{C}^\star$

[18] Indeed, in many ways this is one of central motivations behind much of Carnap's work. To cite just three examples: (Carnap 1969) explicitly states that there are multiple construction systems; (Carnap 1937) goes even further, considering how science is possible when there is not only a plurality of possible languages, but of possible syntactic forms; (Carnap 1950) argues that many classic philosophical debates are nothing but debates over different ways of talking.

is a strictly language-dependent conception of theories: for had there been any difference between $[T]$ and $[T']$ this would have implied that $[T]$ and $[T']$ were different $\mathcal{C}^*$-theories. Now here's the crucial point. Despite being strictly language-dependent, $\mathcal{C}^*$ does *not* individuate theories in an overly fine-grained way. Thus, strict language-dependence does *not* lead a conception of theories to fall afoul of The Individuation Objection.

Of course, the success of this maneuver hinges on the notion of inconsequential variation—a notion likely to raise eyebrows. But here's the thing: To get off the ground, The Individuation Objection presupposes that we have a basic handle on what constitutes an inconsequential variation. For without at least some understanding of which differences are inconsequential and which are not, how can it be maintained that scientific theories can be expressed in different languages? Put differently: The Individuation Objection is, in essence, just the objection that the positivist conception mistakes inconsequential differences of language for substantive differences in theory. So The Individuation Objection is only as good as the present coarse-graining maneuver. Now, it might be objected that The Individuation Objection only requires relatively clear examples of inconsequential variation while the coarse-graining maneuver requires a more substantive understanding of the notion. But for any purported counterexample, the notion of inconsequential variation may be expanded to encompass it. Such particularism may be messy, but it's not clearly objectionable either. Thus, there is a dilemma. On the one hand, if The Individuation Objection gets off the ground, then the notion of an inconsequential difference must be relatively clear. And in this case, the present maneuver allows any strictly language-dependent conception of theories to avoid The Individuation Objection. On the other hand, if the notion of an inconsequential difference is so opaque as to rule out the present maneuver, then The Individuation Objection is equally opaque. Either way, the positivist conception of theories, or a conception close to it, escapes The Individuation Objection.

So if The Individuation Objection really were the central worry, the positivist conception of theories could easily be saved by moving from sets of sentences to sets of sets of sentences. But this gambit will surely fail to impress anyone concerned that the positivist conception is too linguistically oriented. Which goes to show that the central worry with positivism's linguistic orientation must be something else.

1.2 CONFLATION AND ABSTRACTION

Closer to the heart of the matter is what I shall call ***The Conflation Objection***: Roughly, the positivist conception of theories conflates the many legitimate and perhaps even necessary roles that language plays in scientific theorizing—e.g., in formulating and discussing theories—with the *scientific theories-themselves*. But whatever scientific theories-themselves are, the objection goes, they just aren't linguistic entities.

Now, the thing with The Conflation Objection is that it's hard to put one's finger on just what the trouble is supposed to be. It's clear that the objection cannot simply be that identifying theories with something linguistic is problematic because theories aren't linguistic. For absent some reason for thinking that theories aren't linguistic, this simply begs the question against the positivist conception. So what reason is there to think that scientific theories just aren't linguistic entities?

The primary reason comes from an analogy between theories and propositions. For example, recall Giere's claim that the positivist conception goes wrong in looking for "something 'linguistic' to be *the theory itself* …". He goes on to explain the alternative:

> Philosophers long ago introduced the more abstract entities called "statements," or "propositions," to handle just this problem [The Individuation Objection]. A proposition is supposed to be, roughly, *what a sentence asserts*. One can therefore assert the same proposition using different sentences, some in English and some in French, for example.[19]

The thought here is that ***scientific theories-themselves*** are abstract entities analogous to propositions, and like propositions, they must be distinguished from the linguistic entities used to express them, what we might call ***linguistic formulations*** of theories. Linguistic formulations are vehicles that transport bodies of scientific information that are theories-themselves. The charge, therefore, is that positivism conflates the two.

Of course, merely drawing an analogy to propositions only relocates the question of why theories *cannot* be identified with linguistic entities. So we must ask: Why *must* linguistic formulations of theories be distinguished from the theories-themselves? In truth, the answer is unclear. And the more one searches, the more one begins to suspect that the answer is not so much that we *must* distinguish linguistic formulations from theories-themselves as it is that the impulse to differentiate is simply an unspoken dogma.

Focus on the naïve version of the positivist conception of theories, the one that identifies scientific theories with sets of sentences (and not sets of sets of sentences). If any view is at odds with the Abstraction Criterion, the naïve conception is. But is it? The answer turns on how we understand linguistic formulations of theories. On the face of things, to identify linguistic formulations with theories-themselves conflicts with condition (i) of the Abstraction Criterion. But this is too fast. It sweeps an important question under the rug: What's a sentence? Or, more specifically, does 'sentence' mean 'sentence-token' or 'sentence-type'? The unstated assumption is that it is the latter. This is crucial to establishing that every theory-itself has a unique linguistic formulation. For if the naïve view held that linguistic formulations were sets of sentence-*tokens* while theory-themselves were sets

[19] Giere (1988), 84. My emphasis.

of sentence-*types*, then my copy of textbook B would constitute one linguistic formulation of B-mechanics-itself and your copy would constitute another. Which is to say, our textbooks would constitute different linguistic formulations *of the same theory*. So if linguistic formulations are sets of sentence-*tokens*, then the naïve conception does *not* violate condition (i) of the Abstraction Criterion.

Indeed, in this case, any apparent tension between the naïve conception and the Abstraction Criterion dissolves. Since sets of sentence-types are more abstract than sets of sentence-tokens, it follows that theories-themselves are more abstract than their linguistic formulations and so the naïve conception would satisfy condition (ii). Further, the relation between sentence-tokens and -types is plausibly understood as one of expression or assertion. For look at it this way. The orthodox view is, roughly, that the *expression* relation holds between sentence-types and propositions. So understood, the key thing about the *expression* relation is that it partitions sentence-types into equivalence classes and in doing so abstracts away their representationally irrelevant aspects. But here's the thing: The relation that sentence-tokens bear to sentence-types differs little. It too imposes a partition that abstracts away representationally insignificant details. So if there is a difference between the token-to-type and type-to-proposition *expression* relations, it is at most a difference in the relative concreteness of the relata (see figure 4.1). Which is to say, if the naïve conception understands linguistic formulations as sets of sentence-*tokens*, on the one hand, and theories-themselves as sets of sentence-*types*, on the other, then it does not run afoul of condition (iii).

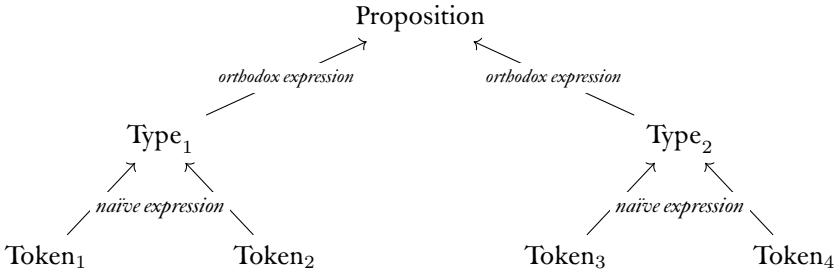


Figure 4.1 *Two conceptions of the expression relation.*

Consequently, the force of The Conflation Objection cannot be that the positivist conception conflates linguistic formulations with theories-themselves while postpositivist conceptions do not. At most, its force can be that the positivist conception fails to push the abstraction far enough. But in essence, this just is The Individuation Objection. For to object that the positivist conception doesn't push the abstraction far enough is the flip side of saying that it incorrectly classifies inconsequential differences between linguistic formulations as ones that signify substantive differences in theory—the correct level of abstraction

is the one where all and only such inconsequential differences are washed away. Yet as we saw, The Individuation Objection, though it weighs decisively against the naïve conception, poses no problem for the more coarse-grained conception that identifies theories with sets of sets of sentences.

The conclusion therefore must be that The Conflation Objection likewise fails to pose a deep problem this coarse-grained conception. At a minimum, the onus falls on positivism's critics to say just how this conception conflates linguistic formulations with theories-themselves in some way that postpositivist conceptions do not. Is it that sets of sets of sentences are not abstract enough? Or that *abstraction* is a cheap substitute for *expression*?

2 WORLDLY FACTS OR LINGUISTIC TRUTHS?

So far we've considered two renderings of The Linguistic Objection and found each wanting. But more importantly, we've found that the notions of a linguistic orientation and of language-dependence are not black-and-white. It is not that some conceptions of science are linguistically oriented while others are not. At most, some conceptions place greater emphasis upon language than others do. So the difference between positivism and its rivals is not some conceptual chasm, as perhaps it first appeared, but instead a difference of degree and emphasis.

How then should we understand the force of The Linguistic Objection? As I see it, the real complaint against the positivist conception—the complaint that has shaped much of the contemporary philosophy of science—has little to do with language, *per se*, but instead concerns the nature of semantic notions. My contention is going to be that what is often presented as an objection to positivism's linguistic orientation really is targeted at positivism's understanding of these semantic notions which, in many eyes, are inextricably tied to scientific antirealism. Exposing this complaint, however, will take some time—much of the remainder of this paper—for it is murky and almost never made explicit.

The complaint begins reasonably enough with the intuition that science is in the business of investigating what *the world* is like, that the objects of scientific inquiry—black holes and biosystems; magnetism, molecules, and mitochondria; quarks and quasars—are *worldly* phenomena, not linguistic ones. And up to a point, this is surely right. The quarks of *Finnegans Wake* are one thing,[20] the quarks of particle physics are obviously another. But this isn't saying much. To furnish any objection to the positivist conception, the intuition must be made more robust. To this end, it is common to flesh out the distinction between worldly and linguistic phenomena with aid of the contrast between *discovery and invention*. For instance, witness how Chakravartty formulates what he calls 'the preeminent question'

[20] Cf., Joyce (1939), 2.4.383.

of scientific metaphysics: "Are kinds *out there*, so to speak, or are they rather *artefacts of convention*, existing only insofar as classificatory practices are brought to bear by creatures such as ourselves?"[21] More generally, linguistic entities are said to be *coined*, *defined*, *invented* and *introduced* and *interpreted*, *stipulated*, and in countless other respects, *up to us*. The realm of language is, in a word, *conventional*. Worldly phenomena, by contrast, are *out there*, waiting to be *discovered*. They are *objective*.

But according according to the complaint, the positivist conception of theories flouts these intuitions. Now, the charge here cannot be that the positivist conception trades scientific theories for linguistic entities, for then present objection would say nothing that The Conflation Objection does not. The charge might instead be understood to be that the positivist conception substitutes linguistic entities, not for theories, but for physical phenomena. But taken at face-value, this is far too strong. The positivists clearly aren't suggesting that worldly phenomena *are* linguistic entities, e.g., that cells and quarks *are* noun-phrases. That's absurd. The way to understand the charge is instead this: On the positivist conception, a great mass of scientific statements that *prima facie* appear to be about the world—that is, statements whose truth depend upon what the world is like—are instead statements whose truth depend solely upon our language. Or to employ a contentious idiom, what appear to be synthetic, *a posteriori* statements are, for positivism, all too often analytic.

Take as an example the claim that spacetime is curved. On its face, this has the hallmarks of a worldly fact. For one thing, its truth seemingly depends upon the physical world's geometric structure: had physical geometry been different, had for instance the interior angles of triangular spatial regions always summed to 180 degrees, then the claim would have been false. Yet as it happens, physical triangles do not always possess such interior angles, so spacetime is not flat but curved. Further, we tend to feel that the realization of this fact was a significant scientific discovery. But for positivism, the charge goes, this is not so. When the positivist says that spacetime is curved he offers no indication of what the world is like, only of his geometric preferences. The difference between flatness and curvature is therefore merely one of linguistic convention. So although we talk today as if spacetime is curved, we could just as well talk as if spacetime were flat. That spacetime is curved is thus no worldly fact.

Now, one can tolerate such linguistic usurpation of worldly fact in small doses. The number of planets in our solar system is, on its face, a worldly matter *par excellence*. But upon realizing that the question turns upon whether Pluto is a planet or the dwarf of one, we tend to feel that the question is, at least in part, a linguistic matter, one of definition. Yet this is hardly worrying. So, at least on a small scale, linguistic usurpation of worldly fact is no cause for concern. But when widespread and systematic, it takes an ominous turn.

[21] Chakravartty (2011), 158. My emphasis.

Just how much of the worldly domain may be annexed by language before turning toxic is, of course, a vague matter. What is not, however, is the usurpation's dark and disastrous denouement. For when the whole realm of worldly fact has been subsumed by language, 'facts' about hadrons, genes, and the like, cease to differ from the myths of Homer's gods in all but degree.[22] But at least in many cases, we feel that this must surely be wrong. So if there is a serious charge, it is that the positivist conception is guilty of *large-scale* linguistic usurpation: that the overwhelming bulk of *prima facie* worldly facts are, given the positivist conception of theories, merely true in virtue of our language.

Yet while this may sit poorly with our intuitions, one may fairly ask: So what? What exactly is the problem? There are perhaps any number of answers, so allow me to skip ahead to the most compelling: that large-scale linguistic usurpation leads to ***conventionalism*** and, in turn, ***scientific antirealism***. The basic thought is that if scientific 'facts' are true in virtue of language, then the truth of such 'facts' is, in some fairly straightforward sense, up to us and our conventions. And this, to the lights of many, is part and parcel of scientific antirealism. As Chakravartty explains:

> This recourse to convention is the primary spur to antirealist rejections of the metaphysical dimension of realism. The strong implicit (and sometimes explicit) *commitment of many realists to a rejection of any such role for convention creates the impression that pluralism is of necessity incompatible with realism*, and thereby, of necessity, associated with some or other form of antirealism.[23]

(Of course, one may wonder what's so bad about scientific antirealism, but for the sake of argument let's just assume that antirealism is unattractive.) The question then is: How does the positivist conception of theories bring about conventionalism and, in turn, scientific antirealism?

2.1 CORRESPONDENCE AND CONVENTIONALISM

The belief is apparently widespread that conventionalism and antirealism are the byproduct of the positivist conception's notion of ***correspondence rules***. The rough picture of correspondence rules, if you recall, was of rules for translating between theory and observation. They are, in a manner of speaking, conduits through which theoretical elements imbibe empirical significance. Yet if this was all there was to correspondence rules, if they were nothing but rules of translation, then they'd hardly be controversial. For even if one takes issue with their adequacy as an answer to the question of what separates mere formalism from full-blooded scientific theory, it's hard to see the question itself as spurious:

[22] Cf., Quine (1951). *Nota bene*: Quine is somewhat ironically lodging the opposite complaint against positivism, viz., that positivism sees a principled difference between physical phenomena on the one hand, and the gods of Homer, on the other.

[23] Chakravartty (2011), 163–164. My emphasis.

everyone has to admit that orthography alone does not explain the relevant difference between "$F = ma$" and "$X = yz$".

But correspondence rules are not just a proposal about the ties between formalism and observation. They comprise, rather, a much more general proposal about the relation of Theory and World, with capitals T and W. And it is here that correspondence rules become contentious. The proposal is not just that overtly mathematical symbols and equations must be tied to observation, but that the whole realm of Theory must be. In this respect, the sentence "Force equals mass times acceleration" stands on a par with "$F = ma$": absent correspondence rules, the terms of each are nothing but empty symbols. For instance, Hempel writes:

> A scientific theory might therefore be likened to a complex spatial network: Its terms are represented by the knots, while the threads connection the latter correspond, in part, to the definitions and, in part, to the fundamental and derivative hypotheses included in the theory. *The whole system floats, as it were, above the plane of observation and is anchored to it by the rules of interpretation* [i.e., correspondence rules]. These might be viewed as string which are not part of the network but link certain parts of the latter with specific places in the plane of observation. By virtue of those interpretive connections, the network can function as a scientific theory: From certain observational data, we may ascend, via an interpretive string, to some point in the theoretical network, thence proceed, via definitions and hypotheses, to other points, from which another interpretive string permits a descent to the plane of observation.[24]

The insinuation being that without correspondence rules or some other suitable connection, Theory floats free of World—that, on its own, Theory is contentless.[25]

It's not hard to see how this leads to a trivial sort of conventionalism.[26] As an illustration, imagine a theory T whose theoretical language L_θ consists solely of a monadic predicate 'gimble' and a constant symbol 'slithy toves'.[27] Since L_θ is a purely formal calculus, the sentence 'slithy toves gimble' could mean just about anything: that grass is green, that snow is white, or any number of other things. It all depends on how T's correspondence rules pin L_θ to its observation language. But this is an arbitrary matter: 'grass is green' is no better or worse as an interpretation of 'slithy toves gimble' than 'snow is white' or countless other sentences. So if T's correspondence rules anchor 'slithy toves' to grass rather than snow, they do so only as a matter of convention. Therefore, it is, in some sense or other, partly a matter of convention whether 'slithy toves gimble' is true and hence whether the world is as T says it is. Such trivial conventionalism points only to the fact that we are the masters of our

[24] Hempel (1952), 36.

[25] 'Contentless' here means *sinnlos* or lacking roughly what Frege called a sense, rather than *unsinnig* or being nonsensical.

[26] Cf., Carnap (1937), section 62.

[27] See Carroll et al. (1990), 195–196.

own words, that we can choose to mean by them what we want. Had we used 'snow' to refer to grass then we could have truly said, "Snow is green". But so what? This certainly does not show that the color of snow—or anything else for that matter—is conventional in any interesting sense, in any sense that challenges scientific realism. Who aside from perhaps Socrates in the *Cratylus* would take issue with any of this? So trivial conventionalism is no reason to disparage correspondence rules.

A more radical form of conventionalism results from *re*interpretation. For if we are free to reinterpret theoretical terms as we see fit, then any consistent theory can be made true (or false for that matter) just by our conventional choices. Does the ether exist? Sure! That is, at least if you take 'the ether' to denote, say, my laptop or your coffee cup. Is vitalism a viable biological theory? Absolutely! We just have to first strip all talk of 'the *élan vital*' of its usual significance. Indeed, with enough exegetical ingenuity, any theory can be rendered true by arithmetic.[28] So if conventionalism via reinterpretation is tolerated, then there are few, if any, worldly requirements for truth. And this radical antirealism of a sort.

Why might we think that such conventionalism results from correspondence rules? The answer is, in part, the same as it was in the case of trivial conventionalism: there are few constraints on how signs and symbols of formal calculi can be correlated with worldly phenomena. But whereas trivial conventionalism restricts interpretation to uninterpreted signs and symbols, radical conventionalism goes further, extending interpretation to all aspects of Theory—in a way, radical conventionalism is what its trivial cousin becomes in the limit. And since all of Theory is, on the positivist conception, just a bunch of signs and symbols, all of Theory must be pinned down by correspondence rules. Thus, when combined with the positivists' view of Theory, the trivial conventionalism elicited by correspondence rules turns radical. Yet we can already see that, in a way, this overstates the case against correspondence rules. The positivists aren't suggesting that correspondence rules are required for all aspects of scientific theories, but only—to put it clumsily—for the theoretical-side of theories. In other words, the positivists do not think that everything is up for interpretation: the language of observation, in particular, is assumed to be antecedently understood and fixed.

But not everything needs to be up for interpretation for the charge of redefinition to stick. It only requires that correspondence rules redefine independently meaningful terms. So the question is if, as Hilary Putnam suggests, "theoretical terms are intelligible, in any ordinary sense of that term".[29] Many critics of positivism have certainly thought so. Putnam, for instance, writes:

[28] See Quine (1964).
[29] Putnam (1965), 260.

> In point of fact, a term is "meaningful"—i.e., has meaning in the language—if it belongs to the common language or has been explained by means of terms already in the common language.…'Radio star', 'virus', and 'elementary particle' *are perfectly good (meaningful) expressions in English. If you have doubts, consult your nearest dictionary!*[30]

And this is certainly correct, at least up to a point. There can be no doubt that a good number of theoretical terms are meaningful: 'radio star', 'virus', and 'elementary particle' are indeed perfectly good expressions. Nor can it be denied that the scientific lexicon overlaps with that of everyday discourse, at least to some degree. After all, average Americans and astrophysicists alike all apprehend 'asteroid'. Biologists and bureaucrats both believe, "Buffalo beget buffalo." Chemists and computer consultants concur: "Combustion creates carbon." And…well, you get the idea. But here's the thing. Even if we grant that correspondence rules are unnecessary in such cases, the great mass of theoretical terms fall outside the sphere of everyday discourse. Terms like 'chemiosmotic mechanism' and 'relativistic spacetime', 'DNA polymerase' and 'differentiable manifold', simply are not meaningful expressions of ordinary English. If you have doubts, consult your nearest dictionary! In many cases, then, correspondence rules do not redefine independently meaningful aspects of Theory but rather are all that separate arbitrary signs and symbols from physically significant devices of representation. Correspondence rules therefore do not lead to radical conventionalism by redefinition.

So if correspondence rules lead to some form of conventionalism concomitant with scientific antirealism, it must be a conventionalism that falls between the trivial and radical ends of the spectrum. Now, I can think of a number of views that occupy this space, but allow me to skip ahead to the most promising two, what I shall call ***Nonliteralism*** and ***Projectivism***. Nonliteralism is the thesis that scientific theories, though truth apt, are true or false only when interpreted in some nonliteral way. Projectivism is a more radical yet vaguer thesis: roughly, that the world lacks any structure independent from how we talk and theorize about it. We shall shortly consider what these two views come to in more detail, but for the moment, let's focus on one key point of agreement between them that bridges the gap between correspondence rules and antirealism. Namely, both forms of conventionalism reject The Noncontradiction Principle:

The Noncontradiction Principle. If two theories contradict one another, at most one can be true.

It's not hard to see why one might think that the denial of The Noncontradiction Principle goes hand-in-hand with antirealism, so the important question is: Why should we think that correspondence rules lead to its denial?

[30] Putnam (1965), 259. My emphasis.

Some brief history is in order.[31] Prior to the nineteenth century, Euclidean geometry had been thought to be *the* correct geometry. But to many geometers, there was something inelegant about Euclid's system: namely, that it just assumed the parallel postulate—roughly,[32] that for any line ℓ and any point p not lying on ℓ, there is one and only one line that passes through p without intersecting ℓ. For nearly two millennia, geometers unsuccessfully attempted to derive the parallel postulate from the other axioms of Euclidean geometry. Eventually, it was shown that they couldn't be derived when, *circa* 1823, Nikolai Lobachevsky and János Bolyai independently discovered what we know today as 'hyperbolic geometry'. In it, all of the axioms of Euclidean geometry held, except for the parallel postulate.

This posed a seeming epistemological problem: How do we *determine* or *know* what the geometry of space is? Prior to 1823, there was an *a priori* reason to think that space was Euclidean: namely, it was thought to be the only possible geometry. But after 1823,[33] it seemed that there was no *a priori* reason for choosing one geometry over the other. This led many, most notably Gauss and Helmholtz, to attempt to empirically determine the geometry of space. Their efforts were, suffice it to say, unsuccessful. As the nineteenth century came to a close, Henri Poincaré argued that the trouble with these efforts was that they viewed the question as an *epistemological* one rather than recognizing it was, in fact, a *semantic* one.[34]

To argue for this, Poincaré asks us to consider a world enclosed in disk.[35] In this world, the temperature varies in a uniform fashion, going from hot at the center to absolute zero at the edges—if R is the radius of the disk and r is the distance of a point to the center, then the absolute temperature at r is proportional to $R^2 - r^2$. Moreover, physical bodies in the world shrink in proportion to their absolute temperature so that the colder a body is, the smaller it becomes. Now imagine that the inhabitants of this world took a measuring rod of length L and, starting at the center, walked in straight line towards the edge of the disk making measurements as they went with their shrinking measuring rod. After some distance r they stop and walk in a circle D until they come back to r. But because the measuring rod neither shrinks nor expands on the journey around D, when they tally up their measurements they find that $D > 2\pi r$. And so they take themselves to be in a hyperbolic space!

So far, this sounds like run-of-the-mill *epistemic* underdetermination, but Poincaré offers a different diagnosis: the inhabitants of the disc world aren't being deceived, they just

[31] For excellent discussions of this history and its influence on the positivists, I highly recommend Coffa (1991), chapters 1, 3, 10, and 17; Friedman (1999), chapters 3 and 4; and Ben-Menahem (2006), chapters 2 and 4.

[32] This is actually not the parallel postulate, but an equivalent one known as 'Playfair's axiom'.

[33] And especially after 1868, when Eugenio Beltrami proved that hyperbolic and Euclidean geometry are consistent relative to one another.

[34] Ben-Menahem (2006) makes a very strong case for reading Poincaré in this way.

[35] Poincaré (2001), 55-56.

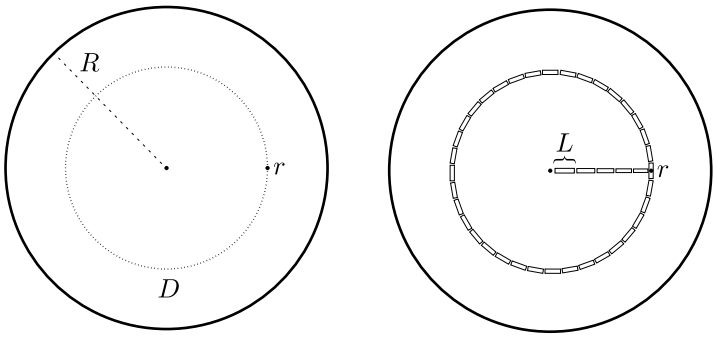


Figure 4.2 *Poincaré's disc world.*

mean something different by geometric terms than we do—they simply coordinate geometric terms such as 'line' and 'circumference' differently than we do. Consequently, "If they construct a geometry, it will not be like ours…*this will be non-Euclidean geometry.* So that beings like ourselves, educated in such a world, will not have the same geometry as ours."[36] While the inhabitants of the world describe it using only geometric terms, as outsiders looking in, we describe it using geometric and dynamical terms. And now here's the crucial thing: our addition of a force compensates exactly for the difference in our geometric descriptions. Thus, as Euclidean geometers, we could "take Lobachevsky's theorems and translate them by the aid of this dictionary, as we would translate a German text with the aid of a German-French dictionary. *We shall then obtain the theorems of ordinary geometry*."[37] So Poincaré concludes:

> *The geometrical axioms are therefore neither synthetic a priori intuitions nor experimental facts.* They are conventions.… In other words, *the axioms of geometry… are only definitions in disguise*. What, then, are we to think of the question: Is Euclidean geometry true? It has no meaning. We might as well ask if the metric system is true, and if the old weights and measures are false; if Cartesian coordinates are true and polar coordinates false. One geometry cannot be more true than another; it can only be more convenient.[38]

In other words, hyperbolic geometry *sans* dynamics is one way of talking, Euclidean geometry with dynamics is another. They are no more in competition with one another than French and English are.

It's easy to see how this apparently conflicts with The Noncontradiction Principle: The inhabitants of the disc world maintain that the interior angles of a triangle sum to less than 180 degrees, while as outsiders we maintain that they always sum to exactly 180 degrees. Our two theories seemingly contradict one another. And yet, if Poincaré is believed, both

[36] Poincaré (2001), 57.
[37] Poincaré (2001), 30.
[38] Poincaré (2001), 45.

theories can be true. This was highly influential in shaping the positivist conception of Theory. For instance, it led Reichenbach to claim: "Thus the geometry of physical space is not an immediate result of experience, *but depends on the choice of the coordinative definition* [i.e., correspondence rule]."[39]

It is this history that lends to the perceived connection between correspondence rules and antirealism. As a result of this perceived connection, few philosophers today want anything to do with correspondence rules.[40] Thus, Giere writes, "Postpositivist philosophers of science have, of course, been *unanimous in rejecting correspondence rules in all their manifestations*."[41]

3 THE TRACTARIAN THESIS

Now, I'm going to argue that the rejection of correspondence rules is the one of the most fundamental differences—if not *the* most—between the positivist and postpositivist conceptions of science. But to see why, we need to first consider what it would be to reject 'correspondence rules in all their manifestations'. And to do this, we need to take a closer look at what it would be to reject Nonliteralism and Projectivism.

Nonliteralism and Projectivism agree about much. The main difference between them is the reasons they give for denying The Noncontradiction Principle. To understand Nonliteralism's reason it will be helpful to contrast it with the trivial and radical forms of conventionalism. Unlike radical conventionalism, Nonliteralism does not see interpretation as entirely unconstrained, but only that it allows for some exegetical wiggle-room, so to speak. Just where is this wiggle-room? It is not just that arbitrary signs and symbols can be interpreted in a plurality of ways—trivial conventionalism recognizes this much. Rather, Nonliteralism goes beyond trivial conventionalism in holding that there is exegetical freedom in how we interpret a theory's *logical form*. For instance, van Fraassen writes:

> On the positivists' interpretation of science, theoretical terms have meaning only through their connection with the observable. Hence they hold that two theories may in fact *say the same thing although in form they contradict each other*… But two theories which contradict each other in such a way can 'really' be saying the same thing only if they are *not literally construed*…

[39] Reichenbach (1958), 19.

[40] This is not to say that the perceived connection was the *only* reason for rejecting correspondence rules. It goes without saying that there were difficulties in ironing out many of the details of how correspondence rules were supposed to work. (For example, see Achinstein (1968), chapter 3.) But, to channel Kuhn, such difficulties with the details only constitute research projects within the positivist paradigm, not reasons to reject the paradigm itself. So there must be some other reason, as Geire claims, to disparage correspondence rules 'in all their manifestations'. And this, as I see it, is their perceived connection to scientific antirealism.

[41] Giere (1988), 75. My emphasis.

> To insist on a literal construal of the language of science is to rule out the construal of a theory as a metaphor or simile, or as intelligible only after it is 'demythologized' *or subjected to some other sort of 'translation' that does not preserve logical form.* If the theory's statements include 'There are electrons', then the theory says that there are electrons. If in addition they include 'Electrons are not planets', then the theory says, in part, that there are entities other than planets.[42]

So look at Nonliteralism this way: If the truth of a scientific theory has three basic aspects—Theory, World, and the theory-to-world relation—then Nonliteralism pins the possibility of true yet contradictory theories on the fact that *there are multiple, equally good theory-to-world relations*.

Projectivism, on the other hand, looks to the World alone for a reason to reject The Noncontradiction Principle. For thing about the World of the Projectivist is that it is, in a manner of speaking, an undifferentiated plenum upon which we project our language, theories, etc. The World of the Projectivist is devoid of object, properties, and the like until we begin theorizing about it. Projectivism is roughly what Jill North has in mind in describing Reichenbach's conventionalism:

> *There is an objective fact about what structure exists, there is a privileged carving of nature at its joints: along the lines of its fundamental physical structure.* Someone like Reichenbach would disagree. Reichenbach thinks that there is *no nonconventional sense to be made of "the" structure of the world*. There is the empirical evidence, and there are different ways of describing that evidence … Neither description is privileged. Neither is a candidate for representing the "true" structure of the world. … *[R]ealism is precisely the denial of the Reichenbachian view.*[43]

So we might say that Projectivism is the view that the World *lacks a privileged structure*. And if the World lacks a privileged structure, then no way of 'carving' the World up is objectively better or worse than any other. Carving the World up into blue and green things, for instance, is objectively no better or worse than carving it up into grue and bleen things.[44] And if this is the case, all that sets one carving apart from others is our conventions. A blue-green-world merely aligns with our conventions while a grue-bleen-world does not. As a result, the Projectivist explains the possibility of contradictory yet true theories by pointing to the fact that the World can be 'carved up' in many different though equally good ways.[45] For example, according to a theory T there are gerrymandered properties, but according

[42] van Fraassen (1980), 10-11. My emphasis.

[43] North (2009), 82. My emphasis.

[44] Cf., Goodman (1983).

[45] Notice that this leads the Projectivist to agree with the Nonliteralist that there are a plurality of theory-to-world relations. For if there is no single worldly structure that a theory must match to be true, then *a fortiori* there can be no privileged theory-to-world relation. But the important point is that Projectivism and Nonliteralism *disagree about the explanation* for this. For the Nonliteralist, that there is no privileged theory-to-world relation goes unexplained. The Projectivist, however, does explain this.

to T' there are not; and yet, for the Projectivist, T and T' may both be true: T is true relative to the grue-bleen-world while T' is true relative to the blue-green-world.

Due to Nonliteralism and Projectivism's incompatibility with The Noncontradiction Principle, many scientific realists have believed it to be imperative to reject both. What does such a rejection involve? The key to rejecting Projectivism is to reject the idea that how we talk and theorize about the World determines its structure. Thus, to reject Projectivism is to endorse the following principle:

(T1) The World has a determinate structure independent of how we talk and theorize about it.[46]

The key to rejecting Nonliteralism is to reject the idea that there are many equally good theory-to-world relations and instead maintain one of these relations is privileged vis-à-vis judging the truth of theories: namely, the theory-to-world relation that preserves the logical form or structure of theories. Thus, to reject Nonliteralism is to endorse the following principle:

(T2) Truth is, at least in part, a matter of matching *the* structure of a theory to *the* structure of the World.

Taken together, (T1) and (T2) form a single principle that, owing to its resemblance to claims of the early Wittgenstein, we might call 'The Tractarian Thesis':[47]

The Tractarian Thesis. A theory T is true only if T has *the same structure as* the World.

And so, to reject both Nonliteralism and Projectivism is to endorse The Tractarian Thesis.

4 RECASTING THE 'LINGUISTIC' OBJECTION

At last, we come to the crux of the matter. My contention is that The Tractarian Thesis is what is really at issue in The 'Linguistic' Objection. And so, more generally, the divide between positivism and postpositivism has less to do with language than with the nature of representation. Of course, it's always dangerous to speculate on what philosophers *really* have in mind when they're objecting to a view, but there is at least strong circumstantial evidence to suggest that this is the case.

For one thing, many postpositivists explain what it is to reject positivism in representational terms. For instance, Suppe states, "abandoning correspondence rules taught us much about how *theories and models attach to nature*."[48] Similarly, Paul Teller writes:

[46] Cf., 'The Independence Principle' in Button (2013), 8.

[47] For instance, "The picture has the logical form of representation in common with what it pictures." Wittgenstein (1961), 2.2.

[48] Suppe (2000), 105. My emphasis.

> An abstract model—a piece of mathematics—can be used by representing agents to represent a target phenomenon by singling out *form or structure shared by the model and its target … Nonetheless it is the model, the abstract object, and the relevant similarity, not the language used to pick them out, which in the first instance function in the representational role.* For that reason among others, an enormous amount of *modern science is deeply extra-linguistic*.[49]

Giere echoes this thought: "*The relationship that does the heavy representational work* is not one of truth between a linguistic entity and a real object, but of *similarity* between two objects, one abstract and one real."[50] And later:

> The links between models are not logical connections, but *relations of similarity*. … The links between models and the real world below are *nothing like correspondence rules* linking terms with things or terms with other terms. Rather, they are again *relations of similarity* between a whole model and some real system.[51]

Such examples are so ubiquitous today that they could be cited *ad nauseam*. Indeed, almost without exception, postpostivists have come to see scientific representation as a matter of similarity or isomorphism.[52] Which is just to say that postpositivist conceptions of science have, across the board, been bound up with The Tractarian Thesis.

Further evidence comes from many of the problems that the positivist conception attempted to address but postpositivist conceptions allegedly sidestep. It is common today to suggest that many of the problems that occupied the positivists were merely the result of their linguistic orientation and so were resolved simply by switching to a nonlinguistic conception of science. For instance, van Fraassen writes:

> Perhaps the worst consequence of [the positivist conception of scientific theories] was the way it focused attention on *philosophically irrelevant technical questions*. It is hard not to conclude that those discussions of axiomatizability in restricted vocabularies, 'theoretical terms', Craig's theorem, 'reduction sentences', 'empirical languages', Ramsey and Carnap sentences, were one and all off the mark—*solutions to purely self-generated problems, and philosophically irrelevant. The main lesson of twentieth-century philosophy of science may well be this: no concept which is essentially language-dependent has any philosophical importance at all.*[53]

But many of these 'purely self-generated problems' are not, as widely believed, the product of language. Rather, they are general problems about the nature of representation. So merely adopting a nonlinguistic conception of science—whatever that is supposed to be—offers no resolution to them. What might resolve them, however, is The Tractarian Thesis

[49] Teller (2008), 436. My emphasis.
[50] Giere (1988), 82. My emphasis.
[51] Giere (1988), 86. My emphasis.
[52] Though there are some exceptions. For example, Suarez (2003).
[53] van Fraassen (1980), 56. My emphasis.

for it is a thesis about the nature of representation. To cite just one example, what elsewhere I call the ***The Problem of Representational Artifacts*** is the problem of saying which features of a linguistic formulation F of a theory T correspond to features of T-itself and which are merely features of F.[54] For instance, does F's typeface correspond to a significant feature of T? This problem was very much seen as one that the positivist conception had trouble dealing with. As Suppe writes:

> Such an approach [the positivist conception of theories] can provide a detailed analysis of the characteristic features of theory formulations, but unless one assume that analogues to these features are distinctive features of theories themselves; *such an analysis reveals nothing about what is characteristic of theories except that their formulations have certain characteristics*.... These considerations suggest that a *direct examination of theories themselves* is likely to result in a more detailed and accurate analysis of the structure of theories than is possible if one concentrates one's attention on theory formulations.[55]

But postpositivist conceptions have been widely believed to sidestep this problem. As Suppe wrote twenty-five years later: "Formalism artifacts that plagued positivism have not occurred."[56] Now, it might be suggested—as Suppe explicitly does—that The Problem of Representational Artifacts was solved by coming to view theories as abstract entities rather than linguistic ones. But on its own, this is no solution. Why couldn't abstract entities also have representationally insignificant properties? In fact, surely they do. For example, *being abstract* and *lacking causal powers* are properties of abstract entities that are insignificant from a representational point of view. So merely viewing theories as abstract entities affords no solution to The Problem of Representational Artifacts. The work is instead done by The Tractarian Thesis. For if it is true, then the question of representational significance reduces to a question about structure.

Though these considerations are admittedly far from conclusive, I submit that The 'Linguistic' Objection has less to do with language than it is an objection to the positivist conception of scientific representation. If this is correct, then The 'Linguistic' Objection goes as follows. If the positivist conception of theories is correct, then correspondence rules are required to anchor Theory to World. But correspondence rules lead to conventionalism of either the Nonliteralist or Projectivist varieties, which, in turn, lead to scientific antirealism. So to avoid antirealism, we must reject correspondence rules 'in all their manifestations' and instead hold that: (i) the World has a determinate structure independent from how we talk and theorize about it; (ii) the truth of a theory is to be judged, at least in part, by whether its structure matches that of the World. In other words, to avoid scientific antirealism we must reject the positivist conception of theories and instead endorse The Tractarian Thesis.

[54] Hirsch (2017a).
[55] Suppe (1974), 221–222.
[56] Suppe (2000), 105.

4.1 THE TROUBLE WITH TRACTARIANISM

The Tractarian Thesis shapes a great deal of the contemporary philosophy of science, perhaps nowhere more so than when it comes to the semantics of science. Semantic concepts like ***theoretical equivalence***, ***entailment***, ***belief***, and so forth, have all come to be understood in structural terms. Indeed, examples are so abundant that they're hard to avoid. The status of The Tractarian Thesis and the conceptions of science it encourages is therefore a crucial question. Is it the case, as Suppe writes in an assessment a quarter-century later, that:

> The [postpositivist conception] has been very successful. It is widely accepted with remarkably little published criticism of it—none fundamental or fatal. It also avoids the problems that scuttled the [positivist conception of theories]: *Correspondence rules are avoided.* … *Formalism artifacts that plagued positivism have not occurred.* It individuates theories better than positivism did.[57]

In closing, I want to suggest that there is a fundamental problem with The Tractarian Thesis and the conceptions of science it leads to.

The key thing about The Tractarian Thesis is that it encourages the view that semantic concepts are unique—e.g., there is a privileged concept of theoretical equivalence, viz., whether two theories have the same structure. Indeed, without this, The Tractarian Thesis offers no solution to any of the problems that allegedly beset positivism. In particular, it would present no true alternative to correspondence rules. The fundamental problem, however, is that there are *many types of structure*. And to each type of structure, there is a different *type of isomorphism*. So The Tractarian Thesis is little better than the thesis that a theory T is true only if it is similar in some respect to the World. But this is obviously vacuous—everything is similar to everything else in some respect or other. So without saying precisely what structure of T must be shared by the World, and precisely what this relation must amount to, The Tractarian Thesis provides no backbone for a theory of representation and the semantics of science.

[57] Suppe (2000), 105.

BIBLIOGRAPHY

Peter Achinstein. *Concepts of Science: A Philosophical Analysis*. Johns Hopkins Press, Baltimore, 1968.

I.J.R. Aitchison and A.J.G. Hey. *Gauge Theories in Particle Physics, Volume I: From Relativistic Quantum Mechanics to QED*. Graduate Student Series in Physics. CRC Press, third edition, 2003a.

I.J.R. Aitchison and A.J.G. Hey. *Gauge Theories in Particle Physics, Volume II: QCD and the Electroweak Theory*. Graduate Student Series in Physics. CRC Press, third edition, 2003b.

Yilmaz Akyildiz. Expository Lectures on Topology, Geometry, and Gauge Theories. In A. O. Barut, editor, *Quantum Theory, Groups, Fields and Particles*. D. Reidel Publishing Company, 1983.

Frank Arntzenius. *Space, Time, & Stuff*. Oxford Univ. Press, 2012.

John C Baez and Javier P Muniain. *Gauge Fields, Knots and Gravity*, volume 6. World Scientific Singapore, 1994.

Jeffrey A. Barrett. On the Nature of Measurement Records in Relativistic Quantum Field Theory. In H. Lyre M. Kuhlman and A. Wayne, editors, *Ontological Aspects of Quantum Field Theory*. World Scientific, Singapore, 2002.

Gordon Belot. Understanding electromagnetism. *British Journal for the Philosophy of Science*, 49(4):531–555, 1998.

Yemima Ben-Menahem. *Conventionalism*. Cambridge University Press, Cambridge, 2006.

George Boolos. To Be Is to Be the Value of a Variable (or to Be Some Values of Some Variables). *Journal of Philosophy*, 81:430–49, 1984.

Phillip Bricker. The Fabric of Space: Intrinsic vs. Extrinsic Distance Relations. *Midwest Studies in Philosophy*, 18(1):271–294, 1993.

Tim Button. *The Limits of Realism*. Oxford University Press, 2013.

Rudolf Carnap. Eigentliche und uneigentliche Begriffe. *Symposion: Philosophi- sche Zeitschrift für Forschung und Aussprache*, 1:355–74, 1927.

Rudolf Carnap. Testability and Meaning. *Philosophy of Science*, 3(4):419–471, 1936.

Rudolf Carnap. *The Logical Syntax of Language*. London, K. Paul, Trench, Trubner & Co., Ltd., 1937.

Rudolf Carnap. *Foundations of Logic and Mathematics*. The University of Chicago press, Chicago, Ill., 1939.

Rudolf Carnap. Empiricism, Semantics and Ontology. *Revue International de Philosophie*, 4: 20–40, 1950. Reprinted in *Meaning and Necessity: A Study in Semantics and Modal Logic*. 2nd edn. Chicago: University of Chicago Press, 1956.

Rudolf Carnap. *The Logical Structure of the World: and, Pseudoproblems in Philosophy*. Open Court, Chicago and La Salle, Ill., 1969.

Rudolf Carnap. *An Introduction to the Philosophy of Science*. Dover Publications, 1995. Reprint of *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*, 1966.

Lewis Carroll, Martin Gardner, and Peter Newell. *More Annotated Alice: Alice's Adventures in Wonderland and Through the Looking-Glass and What Alice Found There*. Random House, New York, 1st ed edition, 1990.

Anjan Chakravartty. Scientific Realism and Ontological Relativity. *The Monist*, 94(2):157–180, 2011.

Alberto Coffa. *The Semantic Tradition From Kant to Carnap: To the Vienna Station*. Cambridge University Press, 1991.

R. W. R. Darling. *Differential Forms and Connections*. Cambridge University Press, Cambridge, 1994.

Shamik Dasgupta. Absolutism vs. comparativism about quantities. In *Oxford studies in metaphysics*. Oxford, 2013.

William Demopoulos and Michael Friedman. Bertrand Russell's the Analysis of Matter: Its Historical Context and Contemporary Interest. *Philosophy of Science*, 52(4):621–639, 1985.

Albert Einstein. On the Electrodynamics of Moving Bodies. 1905.

Albert Einstein. *Albert Einstein, philosopher-scientist*, volume v. 7. Library of Living Philosophers, Evanston, Ill., 1949.

Nancy Forbes and Basil Mahon. *Faraday, Maxwell, and the Electromagnetic Field: How Two Men Revolutionized Physics*. Prometheus Books, New York, N.Y., 2014.

Michael Friedman. *Reconsidering Logical Positivism*. Cambridge University Press, New York, 1999.

Peter Galison. *Image and Logic: A Material Culture of Microphysics*. University of Chicago Press, Chicago, 1997.

Daniel Garber. Old Evidence and Logical Omniscience in Bayesian Confirmation Theory. In *Testing Scientific Theories*, volume X of *Midwest Studies in the Philosophy of Science*. Minneapolis: University of Minnesota Press, 1983.

Thomas A. Garrity. *Electricity and Magnetism for Mathematicians: A Guided Path from Maxwell to Yang-Mills*. Cambridge University Press, 2015.

R. N. Giere. *Explaining Science: A Cognitive Approach*. University of Chicago Press, 1988.

Nelson Goodman. *Fact, Fiction, and Forecast*. Harvard University Press, 1983.

David J Griffiths. *Introduction to Electrodynamics*. Prentice Hall, Upper Saddle River, N.J., 3rd ed edition, 1999.

David J Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, Upper Saddle River, NJ, 2nd edition, 2005.

Richard Healey. *Gauging What's Real*. Oxford University Press, 2007.

Carl Hempel. Problems and Changes in the Empiricist Criterion of Meaning. *11 Rev. Intern. De Philos*, 41(11):41–63, 1950.

Carl Hempel. Fundamentals of Concept Formation in Empirical Science. In *International Encyclopedia of Unified Science*, volume 2. Chicago, 1952.

Marc Henneaux and Claudio Teitelboim. *Quantization of Gauge Systems*. Princeton university press, 1992.

Robert J. Hirsch. Maudlin's Mathematical Maneuver: A Case Study in the Metaphysical Implications of Mathematical Representations. Forthcoming in *Philosophy and Phenomenological Research. DOI: 10.1111/phpr.12226*, 2015.

Robert J. Hirsch. *From Representation to Reality: Interpreting the Languages of Science*. PhD thesis, Princeton University, 2017a.

Robert J. Hirsch. Language, Logical Positivism, and the Semantics of Science. In Erik Curiel and Sebatian Lutz, editors, *The Semantics of Scientific Theories*. 2017b. (Forthcoming.).

Frank Jackson. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press, 1998.

JD Jackson. *Classical Electrodynamics*. Wiley, 1999.

Arthur Jaffe and Edward Witten. Quantum Yang-Mills Theory. In *The millennium prize problems*, pages 129–152. American Mathematical Society, Providence, RI, 2006. URL `http://www.claymath.org/sites/default/files/yangmills.pdf`.

Bence Jones. *The Life and Letters of Faraday*. Longmans, Green and co., London, 1870.

James Joyce. *Finnegans Wake*. Faber and Faber, 1939.

Jeffrey Ketland. Empirical Adequacy and Ramsification. *British Journal for the Philosophy of Science*, 55(2):287–300, 2004.

Erich Kretschmann. Über den physikalischen Sinn der Relativitätspostulat, A Einsteins neue und seine ursprünglische Relativitätstheorie. *Annalender Physik*, 53:575–614, 1917.

M.D. Kruskal. Maximal Extension of the Schwarzschild Metric. *Physical Review*, 119(5): 1743–1745, 1960.

Jeffrey M Lee. *Manifolds and Differential Geometry*, volume v. 107 of *Graduate studies in mathematics*. American Mathematical Society, Providence, R.I., 2009.

G. Lemaître. L'Univers en Expansion. *Annales de la Société Scientifique de Bruxelles*, 1933.

David Lewis. How to Define Theoretical Terms. *Journal of Philosophy*, 67:427–46, 1970.

David Lewis. Radical Interpretation. *Synthese*, 27:331–344, 1974.

David Lewis. Putnam's Paradox. *Australasian Journal of Philosophy*, 62(3):221–236, 1984.

David Lewis. *On the Plurality of Worlds*. Blackwell Publishers, Oxford, 1986.

David Lewis. Ramseyan Humility. In David Braddon-Mitchell and Robert Nola, editors, *Conceptual Analysis and Philosophical Naturalism*, pages 203–222. MIT Press, 2009.

J. David Logan. *A First Course in Differential Equations*. Undergraduate texts in mathematics. Springer, New York, 2nd ed edition, 2011.

George Edward Martin. *The foundations of geometry and the non-Euclidean plane*. Springer, New York, 1996.

Tim Maudlin. Thoroughly Muddled Mctaggart: Or, How to Abuse Gauge Freedom to Create Metaphysical Monostrosities. *Philosophers' Imprint*, 2(4):1–23, 2002a.

Tim Maudlin. *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*. Blackwell Publishers, 2002b.

Tim Maudlin. *The Metaphysics Within Physics*. Oxford University Press, 2007.

Grover Maxwell. The Ontological Status of Theoretical Entities. In Herbert Feigl and Grover Maxwell, editors, *Scientific Explanation, Space, and Time: Minnesota Studies in the Philosophy of Science*, pages 181–192. University of Minnesota Press, 1962.

James Clerk Maxwell. *The Scientific Papers of James Clerk Maxwell*. Cambridge University Press, Cambridge, U.K., 1890.

Charles W Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.

Gregory L. Naber. *Topology, Geometry, and Gauge Fields: Interactions*, volume 141 of *Applied mathematical sciences*. Springer, New York, 2nd edition, 2011.

Mikio Nakahara. *Geometry, Topology, and Physics*. Institute of Physics Publishing, Bristol, 2nd ed edition, 2003.

M. H. A. Newman. Mr. Russell's "Causal Theory of Perception". 37:137–148, 1928.

Jill North. The "Structure" of Physics: A Case Study. *Journal of Philosophy*, 106(2):57–88, 2009.

John D. Norton. General Covariance, Gauge Theories and the Kretschmann Objection. In Katherine Brading and Elena Castellani, editors, *Symmetries in Physics: Philosophical Reflections*, pages 110–123. Cambridge University Press, 2003.

Henri Poincaré. *The value of science: essential writings of Henri Poincaré*. Modern Library science series. Modern Library, New York, modern library pbk. ed edition, 2001.

Hilary Putnam. Craig's Theorem. *Journal of Philosophy*, 62(10):251–260, 1965.

W. V. O. Quine. On What There Is. *Review of Metaphysics*, 2:21–38, 1948.

W. V. O. Quine. Two Dogmas of Empiricism. *Philosophical Review*, 60:20–43, 1951.

W. V. O. Quine. *Word and Object*. 1960.

W. V. O. Quine. Implicit Definition Sustained. *Journal of Philosophy*, 61(2):71–74, 1964.

W. V. O. Quine. *Ontological Relativity and other Essays*. Columbia University Press, New York, 1969.

Hans Reichenbach. *The Philosophy of Space and Time*. Dover Publications, New York, 1958.

Hans Reichenbach. *The Theory of Relativity and a Priori Knowledge*. Berkeley, University of California Press, 1965.

Bertrand Russell. *The Analysis of Matter*. Routledge, London, 1927.

Moritz Schlick, Henk L. Mulder, and Barbara F. B. van de Velde-Schlick. *Philosophical Papers*. Monograph Collection (Matt - Pseudo), 1979.

Raymond A Serway and John W Jewett. *Physics for scientists and engineers*. Thomson Brooks/Cole, Belmont, CA, 7th ed. edition, 2008.

Mauricio Suarez. Scientific Representation: Against Similarity and Isomorphism. *International Studies in the Philosophy of Science*, 17(3):225–244, 2003.

Frederick Suppe. *The Structure of Scientific Theories*, volume 25. Urbana, University of Illinois Press, 1974.

Frederick Suppe. Understanding Scientific Theories: An Assessment of Developments, 1969-1998. *Philosophy of Science*, 67(3):115, 2000.

Paul Teller. Representation in Science. In *The Routledge Companion to Philosophy of Science*, pages 435–441. Routledge, 2008.

Walter E Thirring, Evans M Harrell, et al. *Classical Mathematical Physics: Dynamical Systems and Field Theories*. Springer, 2003.

A. Trautman. Fiber Bundles Gauge Fields and Gravitation. In A. Held, editor, *General Relativity and Gravitation: One hundred years after the birth of Albert Einstein.*, volume 1, pages 287–308, New York, NY, 1980. Plenum Press.

Bas van Fraassen. *The Scientific Image*. Clarendon Library of Logic and Philosophy. Clarendon Press, 1980.

Peter van Inwagen. Meta-ontology. *Erkenntnis*, 48:233–250, 1998.

Robert M Wald. *General Relativity*. University of Chicago Press, Chicago, 1984.

Karl F Warnick and Peter H Russer. Differential forms and electromagnetic field theory. *Progress In Electromagnetics Research*, 148:83–112, 2014.

Eric B Winsberg. *Science in the Age of Computer Simulation*. The University of Chicago Press, Chicago, 2010.

Edward Witten. From superconductors and four-manifolds to weak interactions. *Bulletin of the American Mathematical Society*, 44(3):361–391, 2007.

Ludwig Wittgenstein. *Tractatus Logico-Philosophicus*. Routledge & Kegan Paul, London, 1961. Translation by D. F. Pears and B. F. McGuinness; first published in 1921.

T.T. Wu and C.N. Yang. Concept of nonintegrable phase factors and global formulation of gauge fields. *Physical Review*, 12:3845–3857, 1975.

A Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, Princeton, N.J., 2003.

D. Z. Zhang. C.N. Yang and Contemporary Mathematics. *The Mathematical Intelligencer*, 1993.