Articulating the Core Realist Commitment

by

Nathan D. Morton

Department of Philosophy
Duke University

Date:_____

Approved:

_____
Alex Rosenberg, Supervisor

_____
Kevin Hoover

_____
Andrew Janiak

_____
John T. Roberts

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Philosophy in the Graduate School
of Duke University

2013

ABSTRACT

Articulating the Core Realist Commitment

by

Nathan D. Morton

Department of Philosophy
Duke University

Date:_____

Approved:

_____
Alex Rosenberg, Supervisor

_____
Kevin Hoover

_____
Andrew Janiak

_____
John T. Roberts

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy in the Graduate School of Duke University

2013

Abstract

This thesis comprises an investigation into a very well known and perennial philosophical debate over the interpretive status of our most well confirmed scientific theories, known as "scientific realism." I do not defend scientific realism; rather, I set out to determine what scientific realism is in the first place. My contention is that the thesis is not a single, unified view, but rather a conglomeration of loosely associated propositions that are highly conceptually interwoven, but rarely distinguished. These consist of several different metaphysical, epistemological, and semantic doctrines, which I examine in great detail. I then argue that the indeterminate nature of scientific realism muddles the issue (if there is any) and renders debates fruitless. I attempt to define a thesis with relatively more precise content, which I call the "Core Realist Commitment," CRC. I argue that the CRC prioritizes epistemology – with the thesis that we can and do have (some) theoretical knowledge. I then demonstrate the relatively minimal commitments of the CRC, namely, a minimalist and very undemanding metaphysics, and almost none of the semantic theses that have been traditionally associated with realism. I conclude that the CRC is a step forward in thinking about the debate, not just for its relative precision but also because it is consistent with, and even tolerant of, a wide array of disagreement over concerns that are, I argue, external to the debate and need to be decided on independent grounds.

Dedication

To the memory of Iris Einheuser – teacher, mentor, friend.

Contents

Acknowledgements

I owe gratitude to many people for their contributions and support, without which this thesis would scarcely have been possible. I came to Duke with a passion for the philosophy of science, especially for the realism/anti-realism debate, and Alex Rosenberg helped me sustain that interest until my present views – which unfortunately cannot fully be expressed here – took shape. I am exceptionally grateful for his mentorship and guidance not just on this project, but in general during my time at Duke. I am thankful to Kevin Hoover for his constructive criticism and non-standard views on scientific realism; some of the arguments herein (especially on kinds and mind-independence) are directly due to his gentle nudging of my thoughts in a Peirceian/pragmatist direction (although I didn't quite end up there!). I am also grateful to Andrew Janiak for his helpful comments on earlier drafts. Iris Einheuser, who was originally on my defense committee, sadly did not live to see this project come to fruition. Her influence shaped much of my evolving perspective on these issues – especially concerning the relation of metaphysics to scientific realism. My thoughts on natural kind-hood as a gradational concept and my argument that the hypothesis is *itself* an empirical matter, not something known *a priori* to scientific investigation, while certainly not attributable to her, is the product of many long discussions with her in a directed reading seminar. I am very personally grateful to Karen Neander, who was a source of support and guidance, and who helped me with technical matters and with

advice numerous times. I am thankful to John Roberts, whose seminar on scientific realism at UNC helped further sharpen my focus on scientific realism – especially, at this time, on the difference between global and local arguments for/against scientific realism, and I am grateful for his willingness to sit on my committee on short notice. Thanks also to Stathis Psillos for his kind words, willingness to discuss scientific realism me via email exchange, and for sharing a preprint chapter of his book at a time when I didn't have the money to afford it!

Personally, I thank my mother for many kind words of encouragement when I was closest to giving up or starting over. I owe my father more gratitude for more than I can say in a few sentences – from first giving me Thomas Kuhn's *Structure* to read over the summer – which was the impetus behind my ultimately writing a dissertation on scientific realism in the first place, to helping me with the formatting of the final draft, Dad, this really never would have happened without you. To Kevin and Sean – thank you guys both for constantly hounding me about when I would be finished – I'm pretty sure it helped! I count myself very lucky to have friends like you. Thanks to Mary Kay, for the final swift kick in the pants, so to speak. To Samantha, the love of my life – thank you for always listening to me, for convincing me that I could do this when I thought I never could, for being willing to *learn* much more philosophy than I'm sure you ever wanted to, in the course of helping me talk through my own arguments. My sincerest apologies go out to you for making you upset over the prospect of an infinity

of concrete possible worlds! I promise that out of all the Sam-counterparts in all those other worlds, I am happiest to have you. You have been at my side from the very beginning of this journey, and I couldn't have asked for a better traveling companion. To Morah, I love you so immensely. You bring me such joy each and every brand new day – I am so proud to be able to call you my daughter. Thank you for being exactly who you are.

Introduction

It hardly needs arguing that before evaluating the any philosophical thesis, we need to get clear on what the thesis *is* and what other philosophical commitments adherence to it might entail. My goal in this thesis is to articulate a *maximally inclusive* definition of scientific realism consisting of a robust, yet minimal core set of philosophical commitments. I say this core commitment is *minimal* because (I contend) assent to scientific realism involves much less than is commonly supposed. Nevertheless, I believe the version of scientific realism I defend is *robust* not only in the sense that it is logically incompatible with anti-realist attitudes towards certain scientific theories, but also reaches far beyond weak versions of scientific realism that give away so much to their opponents that they are hardly distinguishable from scientific anti-realism.[1]

What I propose will be a recognizable version of scientific realism. I believe this vision of scientific realism I articulate is more defensible than what has come before. Much more importantly, in so far as a philosophical thesis has an "essential" or core idea that unifies different versions and expressions of it, my proposal comes close to capturing this idea *and nothing more*.[2] It is, I boldly claim, *the real* thesis at stake; what

---

[1] See, for instance, Stanford (2006), who argues for the near *doxastic equivalence* (equivalent beliefs about the content of scientific theories) between modern "watered-down" approaches to scientific realism and his own Neo-Instrumentalism.

[2] For example, something like the "core" idea of British Empiricism would be the idea that all of our knowledge begins with and is limited to experience with the exception of trivial analytic truths, which exhaust the full extent of our *a priori* knowledge. Things like Locke's representative realist theory of ideas, a distinction between primary/secondary qualities, or Berkeleyan idealism are outgrowths, *add-ons* to this core thesis of agreement.

ultimately matters to settling the debate, whether for or against. We are not free to just redefine a philosophical thesis or words in general however we prefer, once an agreed upon use has been settled upon. The debate emerged in a specific historical context with the goal of defending the rationality, objectivity, and most importantly, *accuracy* of scientific knowledge against positivist (empiricist and instrumentalist) and Kuhnian post-positivist (historical/constructivist) conceptions of knowledge, although, as van Fraassen (1980) has pointed out, the debate extends at least as far back Ossiander's preface to Copernicus' *De Revolutionibus*, and probably further to the critique of Ancient Greek proto-scientific or naturalistic accounts of the world by skeptical schools of thought. Because my proposed version of scientific realism represents the core of what is at stake, I call it the "Core Realist Commitment" (abbreviated, **CRC**).

I believe that the **CRC** is the way forward.[3] Unfortunately, this specific project is not aimed at *defending* scientific realism. A sound defense of the version I propose here would be an interesting and worthwhile project and one I intend to undertake in the future. But *this* is rather an attempt to articulate and defend a new approach to the debate, one that dispenses with much of the old orthodoxy.

In fact, this project began as an attempt to offer a new argument for a new version of scientific realism, the one I will lay out here. In the course of writing the introductory chapter, proposing my new approach to the debate, I discovered what I

[3] I have personal reasons for distinguishing these two notions, which will be elaborated on later.

had to say was a dissertation-length topic worthy of exploration unto itself. Much more *needs* to be said about how we should think of scientific realism *itself* as a philosophical perspective on the tasks and achievements of science. A definition is obviously prior to a defense of any philosophical thesis: we need to know what we are arguing about! So this project is best considered as "conceptual housekeeping," as *groundwork* for a future defense of scientific realism that does not depend on what I consider to be unnecessary and in some cases implausible or indefensible positions. The task of an actual defense will have to wait; there is simply too much of interest to analyze concerning how we have thought about the debate itself, and the various ways in which we have been *arguing about the wrong things*.

This dissertation is split into two parts. The first part is what I label the "descriptive task." Therein I investigate the metaphysical, epistemological, and semantic assumptions that have been taken for granted so far in the debate. I will not take any of these assumptions for granted. I do, however, attempt to explain the content(s) of these propositions as fully as possible, as well as offer the best reasons why one *might* think them necessary to a definition of scientific realism. As a preliminary to the descriptive task I introduce a distinction between two forms of arguing for scientific realism, which I call "global" and "local" approaches, following Callender and Magnus' (2004) distinction between "retail and "wholesale arguments".

The second half of the dissertation tackles what I call "the normative task." It investigates how we *ought* to conceive of the scientific realism debate, what it is we *should* be arguing about. This contains the main argument for a core realist commitment. Here I come to reject much of what is ossified doctrines that perhaps *once* had a sufficient basis for being taken seriously as necessary components of scientific realism for the reasons cited in Part I. I argue for what should be uncontentious but has been relatively unrecognized, that certain presuppositions underlying the debate that have to do with the specific context in which the scientific realism debate first arose, as part of the legacy of post-positivism. To preview: I will contend we should deprioritize versions of scientific realism that emphasize *semantic* questions about reference and truth and (at least bold versions of) *metaphysical* theses that make bold claims about the nature and structure of reality based more on philosophical presuppositions than empirical investigation (I have no problem with the latter, but the former has been historically dominant). As I see it, the questions that will decide the matter, pro or con, are essentially *epistemological*. They concern *the quality of our evidence* that certain scientific theories or specific existential claims provide an accurate description of the *way the world really is*. Crucially, my central question is whether we can prospectively identify features or aspects of our best scientific theories/models of the world that are likely to accurately represent things as they really are, those that are not likewise trustworthy, and those for which "the jury is still out" – those too speculative, underdetermined by current

evidence, weakly supported, or for whatever other reason we should not place our confidence in.

1. Global vs. Local Forms of Realism

There is, technically speaking, a "fourth" dimension we could add to the axes of the scientific realism debate, one that cross-cuts along each of the other axes, in the sense that it applies different ways of thinking about and arguing for scientific realism to each of them, separately. I speak of the distinction between *global* forms of scientific realism and *local* ones.

Global scientific realism is, in its most extreme form, the attitude that *all* contemporary theories that have great explanatory power, are well-confirmed by evidence, and command a widespread consensus from the scientific community are accurate representations of what the world is really like. Global realists typically have some overarching philosophical argument for the reliability of scientific methods in general, which makes little contact with the particulars of each specific science. On the other extreme, local scientific realism is limited to the purview of a particular theory or unobservable posit, based on reasons that are contextualized to the content of the very theory *itself*. One might be a local realist about atoms, for instance, on the particular strength of the consilient evidence for their existence (e.g., Perrin's near exact agreement among a multitude of different experiments) or on the basis of some other argument (perhaps our ability to manipulate them in finely controlled ways), and yet have an anti-realist attitude towards most of our contemporary theories (she might believe there are atoms, but distrust the hypothesis that they are constituted by subatomic quarks, due to

the physical impossibility of detecting a free quark). As the example indicates, there are both global and local anti-realist attitudes towards science as well. Most of those persuaded by a general philosophical argument for anti-realism (say, the social constructivist position that there is only a world-as-we-interpret/construct-it) is global anti-realists; they have some broad reason for distrusting any theory. A local anti-realist, on the other hand, could just be a scientist who isn't convinced that the evidence for a theory or unobservable posit is good enough. Skeptics of the atomic hypothesis in the Nineteenth Century could be considered local anti-realists in this sense (this is not to say that all such skepticism was local; Ernst Mach, for instance, doubted the existence of atoms for quite general philosophical reasons, even as the tide of scientific consensus turned towards atomic realism).

In between the two poles are more moderate versions. On the unobservable side of the coin, local realists are likely to only accept those posits that can be detected either with very simple and well-understood equipment, or with many different kinds of instruments. On the theoretical side, they are likely to only accept low-level theories or even just in the existence of specific causal mechanisms. Moving towards the center, the local realist could countenance unobservables that cannot be detected but are instead inferred on the basis of many well-supported and converging lines of arguments (for example, quarks cannot be "pulled out" of atomic nuclei due to the phenomenon of color confinement, but it is my understanding that their existence is inferred on the basis

of very strong theoretical reasons). She might be willing to accept a particular high-level and unifying theory, such as the atomic hypothesis and the atom's quantum mechanical description in the field of chemistry.

From the other direction, the global realist might moderate her position by hedging the claim that all successful scientific theories warrant a realist interpretation. She might do this in two ways. She could restrict herself to specific disciplines, such that for instance she is a realist when it comes to theories within physics, an anti-realist about economics, and undecided about biology. Or she could hold a realist outlook in general, but be willing to give up realism or even be an anti-realist for specific cases. There are many reasons one might do this. There might be good reason to suspect that whatever institutional or methodological mechanisms normally produce approximately true theories are either not in place or are malfunctioning during the development of some particular theory. A theory could be far too speculative and detached from possible observations (as is string theory in its current state), or its own founders might only advance the theory because as a good predictive model without any representational accuracy (for example, Bohr *knew* his early model of the atom was inconsistent with classical electrodynamics and probably deeply false, but put it forward anyway because it got atomic spectra right, and might therefore have led to a better model). This way of approaching scientific realism might be called *ceteris paribus* realism. In the absence of countervailing reasons, we can safely regard theories as approximately true. At any rate,

I will apply this distinction to each of the ideas I discuss, as it reflects an important set of assumptions behind one's reasons for choosing sides in this debate.

The global/local realism distinction is important, so I will take special care to note where certain theses can be interpreted in either form (in which case I will describe the nature of *both*), or when a thesis is essentially global (true in general, topic-neutral terms). Along the way, I will offer some brief commentary on each thesis and their logical implications.

Finally, I need to explain the notation I will use throughout, and some of my assumptions: when I describe any of the particular (perhaps partial) definitions of scientific realism, I give the particular thesis or proposition a title that reflects its basic content. But I also use abbreviations for the sake of convenience, instead of always writing out the full title. The abbreviations generally conform the following logical schema: $\mathbf{\Phi}(t)$.

'$\mathbf{\Phi}$' denotes a second-order variable. 't' denotes a first-order variable, ranging over individual things. The universe of discourse, or the range of possible substitution-instances for first- and second-order variables, consists of all individual *actual* and *hypothetical* scientific theories for the individual variable t – this includes past, present, and any future theories, and the *properties* of these theories for the second-order variable $\mathbf{\Phi}$. It may seem odd to treat theories as individuals, because theories are typically

understood to be either a *collection* of sentences in a formal language,[1] or, more recently, *sets* of mathematical (or other types of) models (Fraassen, 1980). This is a perfectly valid objection, and all I have to say in response is that, contrary to fact, I am *stipulating* that theories are individuals. I could offer an analysis in higher-order logic but this would complicate the technical details far beyond what is necessary to make my general argument. I recognize that theories are not unanalyzable wholes that can be decomposed into smaller parts – whether these are propositions or subsets of the collection of models identified with the theory, and furthermore that sophisticated versions of scientific realism are more selective in identifying *these* "sub-theories" as the proper objects of realist commitment. But again, since what is at issue here is not an argument for scientific realism, but rather an investigation into the concept itself, modeling this would add unnecessary complexity. Where the "individuation problem" of tailoring one's realistic commitments to sub-theories does arise, I will take care to mention this explicitly.

Additionally, I see no problem with extending the universe of discourse to include sub-theories. This partitioning of theories can be principled, so long as the sub-theories thus formed have an internal, coherent logic to them. As an example, the early quantum theory of the atom contains within it the *sub-theory* that the atom consists of protons, located at the center of the atom, that surrounded by negatively electrons that

[1] I.e., the axioms of the formal system are the laws of the theory, provided we have a natural language interpretation of theoretical terms and predicates.

can only occupy discrete energy states, and that the spatial separation between the protons and electrons is vast relative to the size of these individual particles – that is, atoms are mostly empty space, all of which is still true by our current lights. The greater theory in which this sub-theory is encapsulated also postulates definite trajectories, circular electron orbits, it leaves out the existence of neutrons, etc. – parts of the theory we currently reject.

A sub-theory couldn't be just any gerrymandered collection of logically atomic facts belonging to whatever scientific theory we currently happen to believe, however. The conjunction of the fact that that DNA is the unit of inheritance with the fact that the Higgs field is what determines masses of sub-atomic particles is obviously *not* a sub-theory! A sub-theory must be "held together" by relations of *relevance*, and while relevance is a notoriously ambiguous notion, this is only due to its context-sensitivity. The statements "there is a traffic jam on the freeway" and "my plane leaves for Paris in two hours" are not relevant to each other at all, given the background knowledge that I'm taking the train to the airport, but they are of course relevant if instead I plan on taking the freeway. We understand relevance well enough in ordinary conversational contexts, and though vagueness and borderline cases will inevitably arise, I contend the same is true of theoretical contexts. Here we take all of our background *scientific* (and factual) knowledge into account to determine which facts are relevant and which aren't. Crucial to this are our current beliefs about what the best explanations are for certain

phenomena, and about what sorts of real nomothetic relations exist in the world (whether these are *ceteris paribus* causal mechanisms or strict laws). We might have trouble *formalizing* this notion of "relevance," but our ordinary understanding will do just fine. I openly admit that individuating sub-theories is no easy task, but then again, it is well known that individuating scientific theories proper can be just as difficult – e.g., one particular thorny issue is deciding when two mathematical formalisms are mere notational variants of the same theory, or whether their content *qua* interpretation of the formalism yields two different theories.[2]

Our universe of discourse for second-order variables is restricted to predicates that can be *intelligibly* predicated of theories. A scientific theory cannot be green or disagreeable; nor can it have *any* color other than green any more than it can be agreeable. Technically speaking, **Φ** ranges over all properties that can meaningfully be predicated of a scientific theory – even complex, arbitrary "abundant" ones (Lewis, 1983), such as the property of being doubted by Einstein if and only if it is both accepted by Bohr and rejected by Schrodinger, or the property of being formulated before October 7th, 1964 at 8:41 PM – obviously useless properties for the task of analyzing scientific theories.

---

[2] And there is the related issue of when two theories share a formalism but are entirely different in theoretical and empirical content; for instance, Copenhagen versus Bohmian interpretations of quantum mechanics and the meaning of the wave equation – does it represent *the state of our knowledge*, or a physically real process?

Other properties of scientific theories – simplicity, explanatory power, instrumental value, and the other so-called "theoretical virtues" *are* quite interesting and receive the attention they well deserve within the philosophy of science.[3] But these aren't the properties I'm interested in either. Scientific realism is, if anything, a "meta-theory" of how we *ought* to interpret scientific theories. The *interpretive* questions surrounding scientific theories can generally be located along one of the three aforementioned dimensions: metaphysical, epistemological, and semantic. I will focus my analysis on only these.

So, when I use the symbolic notation $\mathbf{\Phi}(t)$, I am attributing some property $\mathbf{\Phi}$ to some scientific theory or sub-theory, t. From now on, for the sake of simplicity I will not explicitly use the phrase "sub-theory," except for where explicit commentary on the perceived need to restrict one's realistic commitments to only selective components of a theory is called for. Therefore I will implicitly assume that any internally coherent, systematic representation that is also a proper part of a past, present, or future actual or hypothetical scientific theory to be within the universe of discourse for t.

I abbreviate the properties $\mathbf{\Phi}$ with a letter to represent one of these three dimensions I mentioned above, with subscripts to identify particular, specific interpretive questions. The global/local distinction is captured in the fact that each

---

[3] Sometimes, these virtues are appealed to in order to solve the argument from the underdetermination of theory by evidence against realism, in order to break ties between rival theories. So these properties are certainly relevant, but a bare exposition of scientific realism won't attribute properties like simplicity or explanatory power to theories *directly*.

partial definition of scientific realism is still stated as a logical schema or open sentence by leaving ***t*** as a free variable, consequently, each proposition of this logical form is left open (both in the ordinary and logical senses of the term "open"), therefore I manage not to pre-decide the question of exactly *which* theories we ought to be realists about. The same property may be instantiated by multiple, most, nearly all, or even *all* scientific theories. The *strength* of various arguments that would license realistic commitment, the specific requirements that any theory must satisfy for it to be interpreted realistically and therefore the *range* of theories that will fall under the scope of these definitions will vary. Some of the proposals I will consider lend themselves more naturally to a local reading, others strongly to a global reading.

2. The Metaphysical Dimension

"Realism" is a term that gets batted around a lot in metaphysics. We have, for instance, modal realism, realism about abstract objects like numbers or universals, realism about propositional attitudes like desires and beliefs, (most famously) realism about the external world, etc. Irrespective of the specific topic, each "realism" shares a common thread. The label "realism" when used in this philosophical context, as opposed to idealism, constructivism, fictionalism, *et cetera*, is most certainly a metaphysical position[1], dealing with the nature of (certain aspects of) reality. Philosophical realism about $\boldsymbol{x}$ is, *ceteris paribus*, the view that $\boldsymbol{x}$ is not *ontologically dependent* upon human minds: roughly, had there been no humans around to say, believe, or do anything, either as individuals or as collectives, $\boldsymbol{x}$ would still exist and retain most of its properties (save for "properties" such as being what Nathan is currently thinking about, or having been scratched by a knife).

However, there are some complications with the concept of ontological dependence. Some things depend on our thoughts and actions in order to both come

[1] I have to be careful here, for there is a very tight connection between "realism" as a metaphysical thesis and "realism" as a semantic thesis, *if* one accepts some general kind of correspondence theory, whereupon *semantic realism* is (speaking *very* roughly) the view that there are such things as abstract propositions, which are *made true or false* by *facts* or *states of affairs* in the world, their so-called "truth-makers." It can be easy to confuse semantic with metaphysical realism, and because of this we need to be careful to avoid confusing the two. And it should be clear that metaphysical realism is more "basic" or "fundamental" as a thesis about the nature of the world, that semantic realism is a *derivative* thesis – true only if its metaphysical version is true. The facts that make propositions true have to be mind-independent and have real, objective existence in order for that proposition to be similarly true independent of what anyone believes, or what concepts one has, and so on. This seems the correct order of explanation. The converse is *not* true, unless we are willing to swallow a *very, very* bizarre metaphysics: it is not a proposition's truth that makes a particular fact obtain, but the other way around – unless somehow believing or agreeing that something is true *makes* a corresponding fact in the world obtain.

into being and to continue existing. One good example is criminality. There is a clear sense in which someone's status a criminal *is* an objective matter. Criminals are people who have violated man-made laws. However, had we legislated differently, some actual criminals would not have been criminals. In a state of total anarchy, there are no criminals. Our activities and decisions *make* criminals in a certain sense; fix the laws of all countries and the actions of all human beings, and you have fixed the extension of the concept criminal. Notice that mind-dependence is perfectly compatible with objective standards for deciding who is or is not a criminal. So, $\boldsymbol{x}$'s ontologically depending on human activity does not necessarily entail a radical form of subjectivism or relativism about $\boldsymbol{x}$. Given a certain legal system, $\boldsymbol{x}$ is either a criminal or not: it makes no sense to say that relative to the way you see things, $\boldsymbol{x}$ is a criminal, while relative to the way I see things, $\boldsymbol{x}$ is not (regardless of whether you or I disagree over whether $\boldsymbol{x}$'s actions as morally justified even if $\boldsymbol{x}$ technically violated a law, or whether we actually even *know* that $\boldsymbol{x}$ committed an act that is by definition criminal).

We have to be careful to distinguish ontological dependence from *causal* dependence, however. There is a sense in which things can depend on our thoughts, utterances, and actions in a much weaker sense. We causally interact with our environment all the time, shaping it to promote our interests. The surface of the Earth is littered with human artifacts, things that would never have existed had we not *physically made* them. This does not mean, however, that artifacts are ontologically dependent on

us, at least in the *same sense* that criminals are. The crucial distinction is between *causal* and *conceptual* dependence. For an object to be conceptually dependent on human activity, all it takes is for a human being or group of human beings to adhere to certain beliefs or perform certain actions or agree to follow certain conventions. Green slips of paper with the right markings and causal history (being issued by the U.S. Government following the correct procedures) have the property being a legal means of exchange for goods and services (in the U.S. and many other places). Paper money only has this property in virtue of the value *we bestow upon it* by a set of conventions (beliefs, practices) that it has such a degree of value. Once we collectively decide to believe and act *as if* some quantity of money can be fairly exchanged for such-and-such goods and services, that amount of money is *ipso facto* worth these things.

It obviously takes something other than mere belief and social convention to make an artifact, like a hammer. Fashioning a hammer requires agential intervention with the world, in the form of gathering the appropriate materials and physically arranging them in the right way. It is true that hammers would not exist if we decided not to make them, but the sense in which we "make" them is through physical interaction with the pre-existing world. There are two ways in which something can causally depend on us: (**a**) it would have never have existed had we not produced it, and (**b**) it would not continue to exist without our continued causal influence. These overlap and also come apart; many objects we make could continue to exist for a long time

without maintenance, and we can intervene to preserve things that we did not originally make.[2] Call the first kind of causal dependence *original causal dependence* and the second kind *sustaining causal dependence*. Realism in the sense of mind-independence seems consistent with both types. In fact, it seems fair to say that we make epistemic contact with many things via *active* causal intervention in the world, mediated by measuring instruments and procedures, that there is a sense in which everything we scientifically has *some* aspects that causally depend on human beings (e.g., being slightly perturbed by a measurement procedure – a simple ammeter, for instance, measures electric current by having the magnetic field generated around the wire move a ferromagnetic object, slightly dampening the current itself). This strain of realism is clearly opposed to conceptual, rather than causal dependence. Its age old opponent is idealisms of various sorts, all of which express the common idea that the world, or some aspect of it, is at bottom a construction out of our thoughts, beliefs, and practices. For my purposes here, idealism is the view that there are no extra-mental objects or properties and that what exists is just whatever we cognize into existence for a certain class of things (e.g., abstract objects, possible worlds, material objects), up to and including the entire external world. Its most famous contemporary guise is in the form of strong "metaphysical" social constructivism, which I discuss in the next section.

---

[2] Such as the active efforts of the National Park Service to preserve the natural granite cliff dubbed "the Old Man of the Mountain," until its collapse in 2003.

Just as much as the other various metaphysical realisms, many have taken *scientific* realism to assert the mind-independent existence of theoretical entities, and the mind-independent reality of causal mechanisms and laws. This view (and a stronger one, as we shall see) is expressed in the definition of scientific realism provided by Psillos. A good working definition is the following:

> [**M**CI (*t*)] *The Conceptual-Independence Thesis*: The domain of nature described by some theory/theories *t* (whether this/these description is/are right or wrong) is ***conceptually independent*** of human individuals and of human collectivities.

Interpreted globally, the existential quantifier "some" ranges over all or nearly all scientific theories; locally, it ranges over only certain theories, possibly including even a single theory. A scientific realist might tailor their endorsement of conceptual independence to exclude certain sciences depending on the details of what the theory studies. For instance, the methodology and findings of the social sciences are subject to concerns of *reflexivity.* These disciplines (I have in mind fields like individual psychology, sociology, political science, and economics) study human *agents*, who are capable of reacting to the findings and developments of these very disciplines, and they can guide policy decisions that can end up affecting individual behavior. Certain social phenomena that are observed, classified, or intervened on in the social sciences is subject

to change; any generalization is potentially *unstable* due to these very activities of observation, classification, and intervention. The social sciences provide paradigmatic case studies since the objects of inquiry and possibly intervention are other human beings, capable of rational deliberation and altering their behavior in response to the effects of developments in these sciences have on their lives.

Concepts of race, class, and gender, for instance, create self-shaping effects as these permeate our attitudes, norms, and the general cultural milieu. Then human beings self-identify themselves as belonging to one of these categories and unconsciously (or perhaps overtly) adopt stereotypical behaviors associated with being a member of that category, thus creating a classic self-fulfilling prophecy; alternatively, someone could reject ("transgress") all of the traditional categories to create new ones as an alternate means of self-expression.

While I have given a distinctly *causal* explanation (in mentalistic terms, no less) of how these kinds of behaviors could be elicited by our own conceptual constructs on a societal level, notice that this goes beyond mere causal dependence of a category on human activities. Counterfactually, if we didn't have these concepts, or if we had fashioned different ones, then these very same behaviors would not have manifested. Concept-possession enters into the causal story and actually effects changes we see *in the world*.

Distinguishing causal dependence from conceptual dependence is particularly important for scientific realism. We routinely produce in the laboratory objects that *could not exist* under naturally occurring conditions. Very heavy transuranic ("synthetic") elements are an example; these are atoms that are radioactively unstable and decay very quickly. There are no naturally occurring samples of such elements due to this instability. The only samples of these elements are ones *we* produce by smashing together atomic nuclei at very high energies. Clearly, the existence of these elements (at least, their existence in the vicinity of Earth) is, in my terminology, *causally* dependent upon the activities of scientists. What matters for realism is that these elements are not *conceptually dependent* upon scientific activity. The scientific realist will want to say that transuranic elements are brought into existence by laboratory procedures, not by mere theory-construction. The *properties* of transuranic elements are not conceptually dependent on our practices either. Having causally produced a sample of ununoctium, we do not get to "make up" just any value for its atomic number. This is fixed, presumably by natural law, at 118. This would be so even if the entire scientific community believed that its atomic number is 117.

This last point about **M**CI suggests another, bolder metaphysical characterization of scientific realism. The independence of theoretically important *properties* from human cognitive activities suggests that theoretical entities and perhaps certain causal processes have a mind-independent, unalterable *nature* or *essence.* This is closely related to the

idea, popular in contemporary metaphysics, that there are *natural kinds*. Natural kinds are classes of objects that have some important natural properties[3] in common. These common natural properties serve two functions. First, they provide the *identity criteria* for the class (necessary and sufficient conditions for kind membership), and thus constitute the kind's *essential properties*. Second, in conjunction with the laws governing these natural properties, their common possession explains more superficial similarities between members of the kind. A quick example is the element gold: all atoms of gold must have an atomic number of 79, and the fact that all gold atoms have 79 protons explains its more directly observable properties, such as its melting point, its specific heat, and its conductivity.

If we take our current best theories seriously, there are at least some *natural* divisions between kinds of entities that hold in virtue of properties that play the essence-role, or, equivalently, provide the identity criteria for being an instance of the natural kind. Strong examples from the hard physical sciences are structural identity (e.g., molecular structures – the arrangement of different chemical bonds or chirality) and *perfect* resemblance (e.g., the indistinguishability of any two neutrinos). Obviously, *identity* of properties (perfectly intrinsic, or complex-structural) will do the trick in explaining the unity (shared non-essential properties) of a kind, as long as there are laws

[3] See Lewis (1983) for a good working definition of what it is for a property to be "natural" instead of gerrymandered, like being grue.

or causal mechanisms that explain why the intrinsic or structural essence fixes the rest of the non-essential properties of members of that kind.

But other properties can serve this "unifying" role as well, and plausible candidates for essential properties will be determined by our total background theoretical knowledge of the putative kind in question. One might want to argue following LaPorte (2004) that biological species have *relational* and perhaps *historical* essences. Given our background knowledge of Darwinian biology, every organism has a common ancestor. Proponents of this theory argue that biological kinds are individuated by a shared lineage, which resonates with thinking of evolution as an essentially historically *contingent* process. The essence of a species, then, is being descended from a most recent common ancestor, or "nodes" on the tree of life. Decisions about where these nodes are, i.e., how to identify the common ancestor, are guided by our intuitions about where speciation events occur, which in turn is guided by our traditional ideas about which animals belong to which kinds, *and* the *actual history* of a species' ancestry.[4]

So this essence concept is intended to do justice to our pre-theoretic intuitions, while also *correcting* deeply mistaken assumptions about actual lineages. We can be very

[4] La Porte thinks that this need for decision-making evidences the fact that essences are *stipulated*, and not *discovered*, making them ultimately conventional, which I think is to a certain degree true – but then again, I think it is implausible to point to anything other than the intrinsic properties of electrons, or structures of molecules, to explain *why* these form kinds. And while biological classification inevitably requires a much greater degree of convention, some decisions are *much* more natural than others, as even La Porte admits. For instance, if we absolutely wanted to retain the primitive notion that whales are fish, we could define "most common ancestor" in such an unnatural way such that most animal life will belong to the same taxon (species). So even here, considerations of naturalness enter into the picture – we are missing out on something very real, and very important if we use such a crude scheme.

misled by our initial biological classification schemes – for instance, on a cladistic taxonomical scheme, reptiles are considered *polyphlyetic* groups of organisms, since the animals we traditionally identify as "reptiles" do not share a common lineage (i.e., they are unrelated organisms that have more than one most recent common ancestor, but they superficially *appear* similar). This is but one illustration of the *dependence of certain essence-concepts on background theory*. It is only because we have background knowledge of the history of life that we can do this; that we can divide up organisms in a principled way, guided by our understanding of objective historical facts and the evolutionary mechanisms that brought these facts about – facts that ground differences in organisms that certainly exist independently of our interests. Other theories require the same *context-dependent* knowledge: for instance, if psychiatry is ever able to discover the *essences*, the unifying properties, of brain/behavior disorders such as autism and schizophrenia – arguably, as of now, the very form of diagnosis as evidenced in the **DSM IV** is to identify a sufficient number of *superficial* symptoms within different categories, each of which are presumably caused by some "deeper" underlying cause.

Dependence on background theory is not even avoidable in the examples from physics and chemistry: it is not *a priori* true that *exact* similarity of perfectly intrinsic properties of (presumably) fundamental particles and complexes (structured arrangements) of those objects and properties are the *essences* of those fundamental particles and of atoms and molecules built out of them. After all, these things have other

properties that they share *exactly* – e.g., all molecules of water are dipolar, and all tau particles have the same mean lifetime – but these are all *consequences* of their essences, their more fundamental properties and structures, and the laws that make these exactly similar but still much more superficial (i.e., derivative) properties "flow" from these essences.

Natural kinds are to be contrasted with *merely conventional kinds* such as the class of fragile objects. Fragility is an essentially dispositional property; fragile objects are those that *would fracture*, under mechanical stress of sufficient magnitude. The concept of fragility serves a very useful task *for us* in sorting out objects that must be handled with special care, but this is a *pragmatic* distinction that does not track any important commonality in the objects themselves. There is no single microstructural property in virtue of which objects are fragile, which the term 'fragile' names. Different materials with vastly different molecular structures can all be fragile. Since no single microstructural property explains every instance of fragility, there seems to be no reason to think the term 'fragile' tracks a real kind-essence. Rather, the boundary between fragile and non-fragile objects seems supplied by humans, as a matter of our pragmatic interests.

Contrast this with our original example, gold. The shared property of atomic number 79 *does* seem to be an important similarity existing independently of our interests. The boundary between gold objects and everything else seems drawn *by the*

nature of gold itself. The doctrine of natural kinds asserts the objective existence of real kinds in nature, differences between types owing to underlying essential properties. Now, science proceeds by abstracting from the particular case to universal generalizations. Classifications and type differences are very important, I would even say "essential" (no pun intended), to scientific work. If there are categories of objects that are dictated by the natural order of things, then knowing what these kinds are, along with their essential properties, must be an important goal of science. We want to develop classification schemes that organize nature according to real differences "out there."

MCI asserts the (partial) independence of the natural world from human conceptual activity, that is, although we may produce objects and their properties by interacting causally with the world, we do not get to *fully* determine what the world is like merely by theorizing about it. We have seen that the doctrine of natural kinds asserts a *special sort* of conceptual independence. Consistently with **M**CI, we might not bring objects and their properties into being by possessing concepts, speaking a certain language, or engaging in some type of social activity: some things have an extra-mental nature unaltered simply in virtue of possessing certain concepts, following social norms, and other such conceptually constitutive activities. Yet, there might be no *natural* basis for preferring one classification scheme over another (of course, there might remain a *pragmatic* basis for the adoption of some classification schemes). Under this hypothesis

we'd live in a world where *extreme nominalism* is true: mind-independent objects might have all sorts of surface similarities, but no essential nature that could serve as the basis for an objective "natural" classification scheme. Friends of natural kinds who also want to endorse scientific realism should find this troubling, and therefore regard **M**$_{CI}$ as too weak. Stathis Psillos feels the need to go *beyond* mere mind independence and assert that not merely the world but also its *structure* is independent of human minds, as evinced by his definition. The next thesis is modified from his own words:

> [**M**$_{NK}$ (*t*)] *The Natural Kind Hypothesis:* The part of the world that some theory *t* attempts to describe has a unique, definite, and conceptually independent natural kind structure (adapted from Psillos, 1999).

Psillos phrases the natural kind hypothesis as a very global thesis, literally, "*the world* has a mind-independent natural kind structure," and in other places he advances the *axiological* thesis the aim of science as limning this structure. Perhaps Psillos is just not being cautious, and it *could* be the case that at least certain sciences are able to discover objective (perhaps merely conceptually independent) facts about the world that group kinds according to differing conventions – i.e., that there could be no one *unique* ("definite") natural kind structure.

On a more local construal of **MNK**, there could be more than one way to "carve up the world's joints"[5] into kinds, which would have different essences or identity criteria and thus (probably) different extensions. The *scientific* motivation for doing so would be that such a multiplicity of kinds would permit different, theoretically interesting and important explanatory purposes, which might be at cross-purposes with one another. In other words, in certain sciences we would miss out on alternative but equally good ways to group things, and certain causal (or nomological) generalizations if we stuck to just a single conceptual scheme. This seems true of biology, for instance, where we have the *phenotypic species concept*[6], *phylogenetic species concept*[7], *biological species concept*[8], *ecological species concept*[9], *genotypic species concept*[10], and *taxonomic species concept*.[11] What would result is either *pluralism* about natural kinds, or *conventionalism* (which would lead to a kind of *metaphysical eliminativism* about kinds – though, remember, not subjectivism or relativism; there would still exist objective identity criteria).

Believing that the world is highly orderly, such that facts about the world can (or *should*) uniquely determine our classification schemes, is not an all-or-nothing

---

[5] To use the eternal metaphor (Plato).
[6] In a sentence: according to this concept, species are groups of phenotypically similar organisms, i.e., populations of individuals with salient morphological and behavioral similarities.
[7] Phylogenetic species were discussed above: they are determined purely by looking evolutionary lineages: a species is the smallest group of organisms that share a common ancestor.
[8] On this view, species are taken to be groups of interbreeding, reproductively isolated populations.
[9] The ecological species concept says that a species are populations that characteristically occupy a specific ecological niche: where it lives and how it interacts with its habitat, its relations to other organisms in the food web, and so on.
[10] Genotypic species are grouped together according to the degree of similarity in the genomes of individual organisms.
[11] Taxonomic species are simply familiar groupings of animals together into hierarchical kinds (kingdom, phylum, class, order, family, genus, species) based on the past decisions of taxonomists (the most famous example being that of Carl Linnaeus).

proposition. These essences don't have to be very robust or perfectly natural, such as that of electrons, which are qualitatively indistinguishable from one another since they share *every* property in common (save for spatiotemporal location, and even then identity gets muddled in the case of entangled particles, say, in the singlet state). I am simply pointing out that different domains of nature might be more or less "orderly," that is, pre-structured by common possession of (more-or-less[12]) natural properties to differing degrees, making "natural" classification schemes more or less legitimate in their claim to marking *really existing* distinctions in the world. So whether **M**NK is true or not might depend radically on *which* theory we are talking about (and what that theory studies).

Finally, there is a third important sense in which scientific realism can be understood as a metaphysical hypothesis. A very straightforward way of characterizing scientific realism is to say that "theoretical/unobservable entities exist." This is an ontological (and *ipso facto*, metaphysical) thesis, expressing the belief that the ontological posits of science (or at least some of them) are real, and actually do have the properties that the relevant theories attribute to them. This way of expressing scientific realism leaves out the crucial role of laws and causal mechanisms as parts of a scientific theory

---

[12] Lewis, in the same account I cited earlier, stresses the gradational nature of natural properties. Some, like charge, spin, mass, quark color and flavor, are, if our best physical theories are true (or true enough), "perfectly natural," in that each instantiation of a *determinate value* for that property/physical magnitude (e.g., being 1.602176×10^{-19} Coulombs – the charge of an electron) is absolutely similar in every respect. Less "perfect" natural properties that are still scientifically respectable are ontologically derivative, like an object's shade of color, specific heat, or viscosity, which can be completely similar *with respect to the behaviors they manifest* in the objects which possess them, but vary widely *with respect to the microstructural, more fundamental natural properties* upon which they supervene (Lewis, 1983).

(though not part of its *ontology*). But we can add to the ontological thesis the belief in the fact that certain causal mechanisms really are responsible for producing the effects attributed to them by scientific theories, or that nature obeys certain laws.[13] Then we get:

> [**M**FT (*t*)] *The Factuality of Scientific Theory Thesis:* The world really is the way theory *t* says it is; that is, the entities that it posits *actually exist*, its theorized causal processes *really are* responsible for producing the effects that are predicted by the theory, and that the "course of nature" *always conforms* to the laws – if any – of *t*.

I have worded the Factuality of Scientific Theory Thesis in such a way as to be amenable to both (rather) global and local readings. An extreme global interpretation of **M**FT is the extremely implausible view that all of our scientific theories have "got it" pretty much exactly right. It is hard to even make sense of such an extreme position, since many current theories are very speculative, weakly supported, or face competing rivals and therefore have not won out in the tribunal of scientific consensus. And, of course, there is the notorious problem that even our best two theories (Quantum Field

---

[13] Notice how carefully we have to approach this: to say that the laws "are true" or "correctly describe the behavior of the world" is not a metaphysical thesis, but a *semantic* one, as the most natural way to interpret these types of assertions is to treat them as attributing truth-values to *law-statements*. Laws *per se* do not *describe* how the world behaves; laws *just are* general facts about how the world always behaves, even under counterfactual suppositions (minimally speaking; I am not endorsing nor offering any particular philosophical theory of lawhood). A statement of these facts, a "law" in the sense of a *law statement* is true iff it accurately represents the special kind of universal fact about the world, the law itself.

Theory and General Relativity) produce both mathematical and conceptual incoherencies if we assume that each theory is complete – a resolution (a theory of quantum gravity) is the remaining "Holy Grail" of physics.

The factual content of the contemporary scientific corpus forbids such an across-the-board global reading. So, anyone who endorses **M**$_{\text{FT}}$ has to embrace *some* nuance, that is, be a localist to at least some small degree: i.e., she must specify exactly which kinds of theoretical entities, causal principles, and laws she believes to be part of our world, and to also articulate whatever qualifications in these endorsements are necessary due to the simple fact that our scientific picture of the world is continually evolving and fallible; practically all theories face some unsolved questions, anomalies, internal and external inconsistencies or interpretive problems, and so on (although I don't think this establishes the *unreliability* of most of what these theories say, at least not without a good argument). For instance, if someone trusts that the world acts very much the way Quantum Field Theory and General Relativity says it does, she would do well to understand where and how the theories conflict and what reality might have to "really" be like according to a more complete account (i.e., the problems a theory of quantum gravity must solve and the sorts of constraints imposed upon it), so as to know where to be cautious or skeptical, and what to expect future change to look like. So if she is rational and well informed, she ought to believe that the ontology and nomological structures of the two theories really are *there*, but our understanding of them is

incomplete (and she may even be able to say why this is so, and how they are incomplete).

MFT is logically weaker *and* stronger than **M**CI and **M**NK in different senses. It is obviously stronger in that it endorses the content of *real*, actual scientific theories whereas the Conceptual Independence Thesis and Natural Kind Hypothesis (**M**CI, **M**NK, respectively) *don't* state or imply that we *actually (or possibly could) have* any grasp of the conceptually independent, *real* objects of scientific inquiry (whether these are natural kinds or not). With respect to the idea that the external world is (mostly?) conceptually dependent on the activity of human beings, the Factuality of Scientific Theory Thesis is absolutely silent. Furthermore, it is *consistent* with a strong form of conceptual dependence. As the idea that there is a conceptually independent natural kind structure to the universe is a logically stronger version of this thesis (i.e., **M**NK (*t*) ⊃ **M**CI (*t*)), the Factuality of Scientific Theory Thesis could easily be true even if **M**CI (the Conceptual Dependence Thesis) is false. Allow me to stipulate that the conceptual dependence of the objects of scientific inquiry on what beliefs and/or conventions that human beings adopt is *logically coherent* – a not untestable claim.[14] Then theoretical ontology could still exist and have a determinate nature, and causes and laws could still have a determinate structure, *even though this determination is fixed by certain human beliefs and practices* (I'll

[14] Kukla (2000) offers convincing reasons to think that total conceptual independence of the entire external world on human beings is inconsistent. It is less clear to me whether the world unveiled by science *cannot* be conceptually dependent *to some degree*, but if the fundamental structure of the world is somehow conceptually dependent on the world, this does seem to create logical problems that seem insoluble.

speculate more on how this could be true later on). My point is simply that acceptance of the theoretical ontology, causes, and laws of (some) scientific theories exist and describe the world we live in need not logically imply that these things, or at least some of these things, are conceptually dependent. (In fact, I think consideration of this possibility *is* conceivable in a certain sense, and that perhaps it is even plausible in a limited domain of inquiry – but again, more on this soon). And, as I mentioned earlier, conceptual independence of what science seeks to discover need not imply that what we believe ourselves to have discovered is totally wrong and off-track: all it means is that *if* science does unveil ontological, causal, and nomological facts about the world, then these facts are conceptually independent.

Now, scientific realism has been taken to involve a lot more, very heady, and rather (in my view) esoteric metaphysical theses. For instance, take Brian Ellis' recent defense of *scientific essentialism* (Ellis, 2001). Scientific essentialism applies Kripke's insights about the rigid designation natural kind terms to scientific kinds. Ellis contends that scientific kinds like electrons all have a particular essence, like possessing the elementary unit of charge or spin ½. Because these essential properties are metaphysically necessary features of electrons, any electron in another possible world *must* have these same exact properties. There's no clear sense in which electrons could have a mass or charge slightly more or less than their actual values. Or, consider Aronson, Harre, and Way's (1995) idea that nature is arranged into a "type hierarchy,"

again according to the natural essential properties of objects. Science succeeds to the extent it maps out the hierarchy. There is the view towards which I am quite sympathetic, Chakravarrty's (2007) *semi-realism*, which posits a sophisticated metaphysical framework in order to motivate positive belief in the reality of certain theoretical *properties*, which are (again) the natural kind essences of theoretical entities and processes, which "cluster" into things (I read this as an endorsement of a "bundle theory" of objects). Chakravarrty, however, stresses that these kinds can be far more conventional than Ellis, Aronson, Harre, and Way think.

These are just three examples; many who have claimed the mantle of scientific realism have also articulated elaborate metaphysical schemes. These are far too numerous to list here, nor should this be much of a worry since we are concerned with the mainstream views with which *most* realists agree – I am trying, after all, to find a minimal "core" view that gets away from what I see as these optional "add-ons" to the thesis. It is certainly clear that, as with any view that involves ontology and laws, philosophers cannot resist getting their hands dirty with very specific metaphysical details. I feel that I can safely ignore these in favor of the "big questions" that every realist seems to worry about – namely, the issues I just raised.

I need to mention that there is also a worry about whether we should be realists about *only* the "fundamental" domains of nature, which would presumably be the entities and laws of fundamental physics. This involves the perennial question of

reductionism vs. anti-reductionism in the philosophy of science, and its ontological implications – i.e., should (or *can*) we be realists about things just like electrons, or do psychological states count, too? Unfortunately I cannot even begin to answer these types of questions here. The issues are just far too complex to receive a fair assessment, given the scope of the rest of the issues I want to address. This is, admittedly, a lacuna in my project, which will require further work. All I will have to say on the matter is that I am an anti-reductionist, even in ontological matters. So I think that we should be scientific realists about things like storm systems or galactic superclusters, *if* there are good reasons to be realists about such things.

3. The Epistemological Dimension

There are many ways in which scientific realism is, and always has been, an epistemic thesis. Scientific realists have always been intensely concerned with the questions "do we have scientific knowledge?" and "is it possible for us to have scientific knowledge?" (and if the answer to either question is affirmative, we should further ask "what *kind* of knowledge?" and maybe "are there good reasons to think that we do *know* these things?").

I will soon argue that in order for scientific anti-realism to even get off the ground, one must identify some reason for treating at least certain types of scientific beliefs as untrustworthy or unattainable. Scientific anti-realism is a form of *selective skepticism* about certain things in the external world – i.e., "ordinary," "everyday," or " non-theoretical" knowledge is secure and immune to most *ordinary* doubt.[1] Most anti-realists are happy to admit that we have genuine knowledge of a privileged sort about the external world. For empiricists these are the "observable" objects of the manifest image of the world, for others (constructivists and neo-instrumentalists), the only knowable facts about the world are non-theoretical, although their distrust of theory can be motivated by a multitude of reasons.

Here, in the context of scientific realism as a debate over the status of our *scientific knowledge*, I would like to briefly discuss some of the anti-realist alternatives.

---

[1] Assuming, of course, that we have some solution to, or some reason to think we can justifiably ignore, the traditional problem of philosophical skepticism.

There are several ways to draw the contrast between realism and anti-realism, but as I see it, the most illuminating way to differentiate the varying positions is to see what they have to say about what we can and cannot know. Like scientific realism, I contend anti-realism is a view about the limits and extent of human knowledge. I cannot *argue* for this here, except to offer a promissory note that this indeed follows from my general argument that epistemology is the issue that needs to be placed front-and-center, since it *motivates* every other proposed form of realism, and since scientific anti-realism is the negation of a primarily epistemological thesis it is, itself, also primarily epistemological in content.

I have already discussed empiricism, especially in its contemporary guise, constructive empiricism, and have found it lacking due to the problem of theoretical-observable entities. The other route for an anti-realist to take is to cast doubt upon the pretensions of theory – to question, perhaps for good reasons, the legitimacy of scientific methods and practices, and the possibility (or mere likelihood) that our theories could come to veridically represent the world.

Historically, the threats to realism from this direction have come from reflection on the dismal track record of certain historically important scientific theories that we now believe to be seriously misleading – the so-called "pessimistic induction" from the past failures of science to the likely substantial falsehood of present theories. This trend was kicked off with Kuhn's *The Structure of Scientific Revolutions* (1970), a seminal but

controversial interpretation of the history of science, which seemed to throw into question the idea of scientific progress or even the idea that science represents a world "out there."[2] Since Kuhn's work seemed to imply that theory-change was non-rational, or at least involved non-rational factors, social constructivists filled in this gap with political, ideological, and social factors – Kuhn's use of the term "revolution" and repeated uses of political analogies at least did nothing to forestall this type of interpretation. This strong form of social constructivism is distinctive from the rest of the views I am considering in that it is, oddly enough, very metaphysical in addition to being epistemological in character. It is the idea, alluded to in the context of **M**CI above, that social processes (persuasion, negotiation, intimidation, and the like) literally *make* the world. And that is a strong form of conceptual dependence "all the way down."

Of course, there is no reason to think that social construction need imply any implausible views about human beings literally "making" the world through a process of social negotiation. A weaker version that is much more intelligible is the view that scientific theories *are* social constructs, but there is little reason to expect that these constructs in any way "match" or "correspond" to the world, even though they are *attempts to represent the world* (this pessimism would stem from the historical facts

---

[2] Though Kuhn himself warned that *Structure* did not imply any such radical relativism and did endorse an idea of historical progress as improvement in scientific "puzzle solving," other of his remarks seemed to indicate otherwise, i.e., the infamous idea that scientists in competing "paradigms" live and work "in different worlds" (150) or that shifts in theory somehow "bring into existence" things like pendulums or the descent of man from primates (this is an avowal of the sort of conceptual dependence I discussed above).

mentioned above, perhaps in addition to other reasons). I take this to be a view implicit in Larry Laudan's critique of "convergent realism" (1981). As this is a non-metaphysical view (scientific facts about the world are not somehow "socially negotiated"), that *constructed* representations are constructs that generally fail to represent the world accurately, whether they are constructed via non-rational processes or not, call this weaker version of social constructivism "epistemic constructivism."

Instrumentalism – the view that theoretical propositions are an uninterpreted mathematical calculus that is useful for making predictions but literally has no meaning – was once a popular alternative to realism. Kyle Stanford has attempted to revive it in recent years, by offering a more plausible version of the pessimistic induction, offering an *mechanism* that serves as warrant for projecting the induction indefinitely into the future – the so-called "problem of unconceived alternatives." (2006) This combines the insights of the pessimistic induction with the *other* general argument for anti-realism, namely, the underdetermination of theory by evidence, the view that there are always genuine rival alternatives to our best theories, and so theory choice must be arbitrary to a great extent.

Stanford argues that the past failure of scientists to think of "scientifically serious" alternative theories to the ones we once sincerely believed, but later rejected as false or seriously misleading gives *present* scientists (and philosophers of science) a good inductive reason to think that there likely are alternatives to our currently preferred

theories, and therefore, at least within the context of "fundamental science," our confidence in our best theories should be undermined – we should not believe ourselves to *know* what they tell us about the world. Stanford's positive view is *neo-instrumentalism*: the view that scientific theories *are* useful for predicting a large number of truths about the ordinary macroscopic world we live in (what Stanford, following Quine, calls "the bodies of common sense"). But we can (of course) still understand what these theories literally say: they have an interpretation, Stanford admits; it is just that the problem of unconceived alternatives to our best theories provides no reason to think that these theories are veridically representational.

The antithesis of these specific versions of anti-realism *affirms* two positive attitudes towards scientific knowledge, one logically stronger than the other. Of course, a full antithesis of anti-realism will encompass much else as well – to take one example, the affirmation of the literal interpretation of the *language* of scientific theories, against the view that they are reducible to their "empirical (observable) content" (reductive empiricism), or that they are uninterpreted calculi (instrumentalism). This would be a semantic version of scientific realism, not an epistemological one. For now, however, we are sticking with epistemology. Both realistic "epistemological attitudes" can be read off Richard Boyd's (1983) definition of scientific realism. Thesis (ii) of Boyd's definition of scientific realism is:

> Scientific theories, interpreted realistically, are confirmable *and in fact often confirmed* as approximately true by ordinary scientific evidence interpreted in accordance with ordinary methodological standards (45).

Boyd asserts two propositions here, and I think it is very important to recognize and separate both. First is the idea that scientific knowledge is *possible*. He says that scientific theories are confirm*able* "as approximately true…by ordinary scientific evidence interpreted in accordance with ordinary methodological standards." That is, in *our* world, were the right conditions to obtain, then we *could* have scientific knowledge (in his words, know that a scientific theory is "approximately true"). And we could do this without having to appeal to extra-scientific standards of knowledge, such as a Cartesian or Kantian (transcendental) form of philosophical justification – we wouldn't need to point to anything *outside* of ordinary scientific practices and findings in order to justify the accuracy of any scientific claims that we *might* happen confirm.[3] The central intuition Boyd is getting at (which he clarifies later on) is that the realist's *possible* stake to possessing scientific knowledge is on par with our more ordinary methods of gathering evidence *in general* about the world. We don't use two separate methods and standard to find out, say, whether a certain suspect committed a murder, compared to finding out

[3] As Marc-Alspector Kelley points out (2001), Boyd's strategy of defending realism by deploying the very popular *explanationist defense of realism* – the idea that approximate truth is the best explanation of the success of scientific theories – *is* an extra-scientific appeal to non-scientific, philosophical intuitions, and thus his argument in the rest of the paper cited belies his own methodological principles!

whether electrons are real. There is just one project of gaining knowledge about the world; there is just *inquiry in general*. True, every knowledge-claim will require some differences of method: different tools, different background knowledge for interpretation of evidence (even when comparing two investigations into "everyday" knowledge) – but the same *standards of evidence* and the same set of *good inferential strategies* are exactly the same, in both ordinary life *and* science. So we have:

> [**E**PP (*t*)] *The Principle of Epistemic Parity:* There is no principled distinction between knowing non-theoretical facts and knowing facts disclosed by a particular theory *t*. It could be the case that the non-theoretical/theoretical distinction does not track *any* significant difference concerning what we can or cannot know, such that any possible knowledge mediated by scientific theories is on an epistemic par with "non-theoretical knowledge" of the macroscopic world around us – the "manifest image" of Sellars (1962). *If* there is an in principle reason why scientific knowledge is typically unattainable, these reasons do not hold in the special case of theory *t*. Conversely, there might be no in-principle significant distinction between non-theoretical and theoretical knowledge as a general rule, but there might be special facts about *another* theory *t** which make its content unknowable. **NB**: This *does not* imply that we *actually have* any such knowledge; only that it is metaphysically possible that we *could* know such things.

I have tried to formulate this principle as *potentially* local. But it is *much more* naturally interpreted in a global sense, as saying that there is simply nothing unique or distinctive about science such that its knowledge-claims are in principle unknowable. I have strongly leaned towards this global interpretation in my phrasing of the Principle – as a thesis about the nature of scientific inquiry in general. This global version, directed in part against at least some philosophical arguments that draw a line in the sand between "ordinary" and scientific/theoretical forms of knowledge. Although there are *selective anti-realisms* about only certain forms of scientific knowledge, and the most plausible arguments for anti-realism are for local versions, there are also global arguments that attempt to undermine the whole body of scientific knowledge in one fell swoop. These arguments work (if they work) by singling out some property of science as a *general form* of inquiry, and not any particular property of certain specific fields of putative scientific knowledge. For instance, the pessimistic induction straightforwardly infers from the near 100% turnover rate of past "scientific"[4] theories to a high likelihood that our present theories are false. Stronger forms of the underdetermination thesis[5] that claim that there are *always* rival theories we can "cook up," even in principle, that are as easily

---

[4] I put the term 'scientific' in scare quotes here only because I (and many other realists) have thought some of Laudan's (1981) examples ludicrous as examples of genuine science, such as the humoral theory of disease, or the astronomical hypothesis of crystalline spheres. Yet, some of his examples (ether theory, phlogiston, and caloric) are admittedly quite damning.

[5] Which Laudan and Leplin (1991) have, quite rightly in my view, refuted.

confirmable as our preferred ones. These arguments don't discriminate among scientific disciplines. They are quite rightly aimed at theoretical knowledge as a *form of inference*, regardless of its *specific content*. Since these arguments would undermine *any* claim to scientific realism, a would-be scientific realist *has* to endorse the Principle of Epistemic Parity in global form.

As the above definition suggests, perhaps the Principle of Epistemic Parity *could* be shown to fail for local reasons that have more to do with tangible worries about the reliability of specific methods and/or inferences that are specific to some particular field(s) of science. Still, the argumentative strategy is the same: every anti-realist argument attempts to burden us with doubts about the reliability of using our normal inferential methods to draw conclusions about domains of nature that are outside of our ordinary experience, including the use of sophisticated instruments to help us "probe" what we cannot otherwise sense. The local anti-realist just points to *content-specific* reasons why these methods can't be trusted in a *particular* domain, but not necessarily across the board.

An example of local, selective anti-realism that appeals to a denial of Epistemic Parity would be Derek Turner's (2007) charge that we can know things about the micro-world, but for all we know, we might actually socially construct the pre-historic past (this is actually a form of *non*-realism; see my remarks below). The argument rests on an epistemic asymmetry thesis: unlike, say, the field of physics, (**a**) our background

knowledge about the past is just too *sparse* to rule out competing explanations, (**b**) worse, our background knowledge of geology, fossilization, etc. actually informs us that physical processes tend to degrade and corrupt the fidelity of information that might be preserved in physical traces, and (**c**) we cannot *manipulate* the past, therefore, we cannot use any experimental methods – the best we can do is try to construct a coherent narrative that fits and best explains all the facts, make a prediction based on this hypothesis, and search for corroborating evidence; we can't *intervene* in the past in the same way we routinely do in physics to *directly* confirm or disconfirm a causal hypothesis. Notice the appeal to content-specific facts about the unreliability of the methods that paleontologists and archaeologists are forced to use – especially the appeal to *other theories* (!) of processes that are likely to degrade and erase traces of the past, or to the fact that the past is epistemically different, in that it is nomologically impossible to interact with. Note also that local realism about the possibility of knowing some theories is of course consistent with local anti-realism about other instances of scientific knowledge. Turner does not explicitly avow realism about physics, but he does suggest it is on a much better footing epistemically, so it is perfectly fine to have specific doubts about the historical sciences, even if one embraces a very generous form of realism that lends credence to most of what all *non-historical* sciences suggest.

Other local, content-specific cases against Epistemic Parity rely even less on abstract considerations like an epistemic asymmetry introduced by the asymmetry of

time (lack of manipulation, degradation of information), but seem to live quite close to the scientific facts themselves, which internally (as a consequence of the *theory itself*) disclose that there is a class of facts about the world that are completely unknowable, undetectable, or both. For instance, from the standpoint of our current state of scientific knowledge, information captured by a black hole is completely erased forever. We *can* infer many things from theory, for instance, that a human who nears the event horizon will deform ("spaghettify") into a string-like shape, and that black holes have *entropy*; they can dissipate over a very long period of time. But as things stand right now, we cannot even physically describe what happens at a singularity; the field equations of general relativity approach infinite curvature near the limit of the point-singularity, and infinite quantities in physics are, generally speaking, an indicator that something is wrong, due to their physical unintelligibility.

Now, of course, no working scientist takes this as a basic fact about the universe; it is, I think, generally expected that whatever resolution of general relativity and quantum mechanics we are able to formulate, a theory of quantum gravity, will smooth out these difficulties. Still, even if (or when) we are able to redescribe a black hole on a quantum level, many of their curious properties are likely to remain, including the unfortunate fact that no signal could escape the event horizon, thus barring any possible direct observation of what goes on inside that boundary. Again, we can *infer* certain facts about what goes on beyond the event horizon on the basis of theory, but it seems that

any kind of possibility of direct verification by experimental interaction is foreclosed. There are other such things that we have (it seems) excellent or at least very good internal theoretical reasons for believing in, but these very same types of reasons also seem to straightforwardly imply their undetectability. Black holes are one example, and to mention just one more, so-called "free quarks," quarks that are unbound into hadrons, such as protons and neutrons, cannot exist in nature according to the Standard Model due to the hypothetical mechanism of color confinement.[6]

So for certain classes of things or certain domains of nature (space-time in the vicinity of the singularity, within the event horizon), it would seem very little extra-scientific reason would be needed for thinking that either (**a**) there are things we cannot know entirely, or (**b**) there are things that we can only know on completely inferential grounds, on the basis of theoretical economy and explanatory power, but for which there is no way for us to causally interact with them. Although, of course, the state of our "knowledge *about* our possible knowledge" is subject to change as our theories are refined (and in both of my examples, there is no reason to expect that what I have said is

---

[6] Quarks are postulated to have "colors" of blue, green, and red (color is like spin and charge in being a fundamental quantum-mechanical property but one that only quarks possess). Other fundamental particles like electrons can exist in free space: according to quantum field theory, the electron creates virtual positron-electron pairs and virtual photons that are responsible for the change in net charge (potential, energy) in the local field. Were a quark able to exist in free space, it would create virtual quark/anti-quark pairs as in the case of electrons, but it would also create virtual gluons (bosons that transmit the strong force). The Standard Model of particle physics predicts that the presence of such virtual gluons existing in free space would induce the creation of more virtual gluons, and a cascading effect would quickly occur: the local strong field would have infinite energy very soon as the amount of virtual gluons would increase without bound. This is a physically absurd situation, and can be rejected *a priori* on the grounds that this would violate the most fundamental known laws of nature. This putatively explains why we don't, and could not ever, see a quark in isolation from the hadrons in which they are bound together.

the end of the story; both of these theories are expected to be able to unified in principle, and therefore altered) – so one might want to take these very local scientific arguments with a grain of salt, and closely examine the internal, theory-specific reasons for why knowledge is *prima facie* impossible to see how insoluble or speculative they really are.

Now, Boyd's (1983) second thesis is to assert the (perhaps much) stronger claim that we actually *are* in such possession of the theoretical knowledge that he indeed believes is possible.

> [**E**TK (*t*)] *The Theoretical Knowledge Thesis:* Scientists *actually know* certain facts about the world in virtue of believing theory *t*. The systematic body of theoretical claims that constitutes *t* represents the world faithfully, as it really is – at least within most important respects and degrees.

Unlike the modal Epistemic Parity Thesis, the Theoretical Knowledge Thesis can have either a global or local interpretation, but extremely global readings (i.e., scientists know pretty much everything that their theories tell them about the world) seem (at least to me) wildly implausible. Some discrimination is probably called for – one could (or perhaps should) believe that due to the interpretive difficulties of the Copenhagen interpretation, only certain aspects of quantum mechanics represent the way things are in the world. For instance, it seems to me that one ought not to construe wave-packet

reduction as a real physical process; that the theory will remain quite unintelligible in the absence of a solution to the "measurement problem." But it also seems reasonable to believe that photons exist, and have properties ascribed to them by quantum theory: spin 1, zero charge and mass; they appear to have both wave- and particle-like characteristics which can be (partially) understood by grasping the Heisenberg uncertainty relations, which help us to understand the fundamental difference between ours and the "fuzziness" of the quantum, that entities appear to be "smeared out" in space and time rather than localized – that is, until they are measured. It is not this unintuitive objective indeterminacy which is *unintelligible* – we can wrap our minds around how the uncertainty relations require that certain properties (position and momentum, energy and time) can only be simultaneously possessed to a certain degree and trade-off one another; what is mysterious is why this indeterminacy does not percolate up to the macro-world: why things existing at our scales of length and energy/mass have classical definiteness, and why measurements on quantum systems "collapse" the wave function and make measurements have definite outcomes (instead of us being able to directly observe a superposition).

Proponents of the Theoretical Knowledge Thesis obviously shoulder a stronger burden of proof. Boyd asserts the *conjunction* of **E**$_{PP}$ and **E**$_{TK}$; theories are both confirm*able* and "in fact often confirmed as approximately true." This is of course redundant: because **P** implies possibly ◊**P**, **E**$_{TK}$ implies **E**$_{PP}$.

As I have made clear, the Theoretical Knowledge Thesis is the crux of the issue of scientific realism; I argue that this is a good candidate for the Core Realist Commitment.

4. The Semantic Dimension

Historically speaking, scientific realism has been characterized in semantic terms, even if this is not commonly recognized as such; perhaps the most familiar (if unrefined) statement of scientific realism to non-specialist philosophers is "scientific theories are (approximately) *true*." Another very popular realist rallying cry is the assertion that "theoretical terms *refer*." Much of the reason for the dominance of analyzing scientific realism in terms of truth and reference owes, in my view, to contingent historical circumstances – specifically, the legacy of the downfall of logical positivism – which I will discuss later, in the broader context of *how* scientific realism became associated with so many different ideas at once. Suffice it to say for now that the verificationist criterion of meaning, coupled with a rather uncompromising empiricism, made the interpretation of theoretical terms into a philosophical puzzle for the logical positivists. Leaving aside the epistemological question of how we could ever *know* about "unobservable" entities we can't experience, there was the semantic question of how "theoretical" terms (recall that the positivists did not distinguish theoreticity and unobservability) could refer or be meaningful at all, especially because the verifiability criterion of meaning would require some possible experience of these theoretical/unobservable entities, which could only be very indirect, via their possible empirical consequences. The two responses to this problem were (**a**) either to dissolve the problem of theoretical terms as a pseudo-question, treating them as meaningless symbols that, when manipulated in accordance

with the postulates (axioms, laws) of the theory and correspondence rules, yielded correct, verifiable, observable predictions or (**b**) reducing their meaning and reference to their "empirical content," treating them as massively elliptical ways of referring to all possible observations.

These attempts to reform or eliminate theory-talk to make the world safe for empiricism met with many obstacles, and had been mostly abandoned by the latter half of the twentieth century. Left without an adequate solution to the problem of explaining the meaning and reference of theoretical ("unobservable") terms, early scientific realists followed their positivist forbearers in retaining a central place for the philosophy of (scientific) language. The ascendant view in philosophy of science was that talk of theoretical entities was to be *taken literally*, on par with ordinary language. Thus the statement "the current in this wire is 200 amperes" was to be understood using *exactly the same* referential semantics that applies to statements such as "the Statue of Liberty is green." A willingness to apply Convention T to theoretical language naturally seemed to imply a literal, and thus realistic, view of scientific theories. So "the current in this wire is 200 amperes" is true if, and only if, the current in this wire is 200 amperes. On its face, this requires that *if* the above sentence is true, then theoretical phenomenon *electric current*, as described by classical electrodynamics as the flow of *charge* (or even quantum mechanically as the "motion" of *electrons*), *exists*, and that the magnitude of the current in this wire actually is 200 amperes. If we are ready to accept any scientific description of

the world as true, as we seem to do *prima facie* when a theory is well-confirmed, then standard referential semantics demands that the things that scientific theories purport to refer to really do exist, and really are the way they are theoretically described (*modulo* worries about idealization, incomplete representation, approximation, etc.).

Because of this close connection – this very intimate tie between what *exists* and what is *true* – ontological realism and truth-conditional referential semantics can tend to blur together if one is not careful to keep them separate. Where e stands for any theoretical entity, and $\{e_1, \ldots, e_n\} \in t$ is a collection of entities are the posit of a theory (that also provides a description of the properties of the e_i's, and their nomothetic interactions with each other: the laws and causal mechanisms that interrelate them), then "what t says about $\{e_1, \ldots, e_n\}$ is approximately true" appears to be no different in content from "e_1's, … , and e_n's all exist, and to an approximate but significant degree, they have the properties and behavior that t says they do." But, assuming one holds to quite popular views concerning truth (i.e., a correspondence theory), these *are* very different statements, and to interpret them as expressing one and the same thing is a seductive mistake. The first sentence is about *truth-bearers*: things that can be true or false, in this case (let us say), a set of *propositions* assigning predicates to, and relations among, the e_i's. Ignoring for convenience the difficult issue of *approximate* truth, the theory (considered as a collection of propositions) is true if and only if the predicates name properties that the e_i's really do have, and the relations that hold among the e_i's

(which are probably nomothetic in character) actually obtain. The "… is (approximately) true" sentence claims that this sort of complex relation obtains between the set of propositions and the world. The specified arrangement of facts in reality is the *truth-maker* for the propositions constituting the theory: the way the world is makes what is *asserted of it* true or false.

The second sentence, on the other hand, just says that the e_i's have certain properties and obey certain laws or stand in certain causal relations to each other, and that is all. It is not about a *further* relation between some intermediary truth-bearer and the world, but purports to say something about the world itself. They are demonstrably different if one just considers their implicit ontological commitments. The second sentence is merely committed to the existence of e_i's, their properties, and nomothetic relations. The first, however, carries an excess commitment to other *entities*, namely, propositions, and the *relations* of truth and reference.

I do not think that many philosophers of science have appreciated this point, although some have (see, for instance, Leeds 2007). For the most part, however, mention of 'truth' has always played a dominant role in statements of scientific realism. Take, for instance, Boyd's defense of realism, in which he asserts that theoretical terms ought to be thought of as (putatively) referential and theories as "literal," or Psillos' updated (1999) defense of there being a robust semantic component to scientific realism, where theories are construed as "truth-conditioned descriptions of their intended domain." Psillos goes

on to argue very explicitly that he intends a correspondence reading of the truth predicate and of the reference relations. Not every scientific realist who endorses the idea that theories are true goes on to explicitly mention that *this* is how they understand the meaning of the word 'truth.' However, up until very recently at least (with the rising popularity of deflationary accounts of truth[1]), the correspondence theory has been overwhelmingly the most popular philosophical theory of truth. Thus, we have:

> [**S**$_{CoC}$ (*t*)] *Conditional Correspondence-Truth Realism*: We assume by default that any correspondence theory of truth when we judge that any particular scientific theory *t* is true (or false). At minimum, correspondence requires that truth is a relation of 'match' or 'agreement' between truth-bearers (representational vehicles, such as propositions) and truth-makers (facts). For there to be such a kind of "correspondence" between assertion and reality, there must be robust relations of world-to-word and word-to-world links, such as meaning and

[1] Deflationism encompasses a very wide range of minority views on the nature of truth, with the shared agreement on two core principles: one, that the truth predicate does not name a metaphysically robust relation between an extra-mental, extra-linguistic abstract proposition that can be expressed by different token utterances and inscriptions in different languages, etc. Second, deflationists make much of the "redundancy property" of truth: the fact that for many (but not all) linguistic assertions, to say the assertion is true essentially adds nothing over and above the assertion; e.g., there is not a clear difference between the sentences "the sky is blue" and "that 'the sky is blue' is true." One theory, disquotationalism, treats the predicate 'is true' as an elliptical way to express assent to whatever sentence(s) truth is being predicated of. In this way the truth predicate still serves a useful function, for instance, if one wants to say, "Everything that Bible says is true" – expressing agreement with a very large set of assertions, without having to assert each one by itself.

reference, and maybe also a kind of structural similarity between truth-bearer and truth-maker.[2]

Now, of course, notice that I didn't say in **S**$_{CoC}$ that any theory actually *is* true, but at the beginning of this section I said that it is commonplace to hear scientific realism described as the thesis that theories *actually are true* (at least approximately so). This is why I have called the above thesis "conditional." Similar to **E**$_{TK}$ (The Theoretical Knowledge Thesis), which was a strengthening of **E**$_{PP}$ (The Principle of Epistemic Parity) by actualizing a modal claim, by affirming the antecedent of the conditional thesis, we derive a logically stronger version of semantic realism:

[**S**$_{CaC}$ (*t*)] *Categorical Correspondence-Truth Realism:* Theory *t actually* ***is*** true in the sense of correspondence described in **S**$_{CoC}$; thus, some theoretical terms genuinely refer to real things in the world, and some theoretical claims about the world (e.g., existence-claims, law statements) are true by virtue of match between the propositions of the theory and facts in the world (the latter making the former *true*).

---

[2] A very simple example: the proposition 'Beau is lying on the couch' is true if and only if there is a structural similarity between the world and the content of the proposition. My dog Beau is lying on the couch. The *term* 'Beau' refers to my dog, and 'the couch' is an indexical term understood by my audience to name *my* couch. I have names for two different objects (each fixed by some relation of reference – naming my dog, just being in the vicinity of a unique couch), to which I further apply the schematic relation "***x*** is lying on ***y***" – substituting 'Beau' for ***x*** and 'the couch' for ***y***. The proposition is now *grammatically structured* in such a way that reflects or parallels the way that the facts in reality I intend to describe are structured.

This statement is arguably the most common, perhaps platitudinous, way of expressing scientific realism, but crucially, with all of the assumptions that are part of any mainstream correspondence theory made very explicit.

Why do most realists insist upon characterizing their views in the language of truth, and especially insist upon the correspondence theory of truth? There are several reasons. First, scientific realists owe us an explanation of what they mean by the "literal" truth of theories. One way to do this without adopting a full-blown *theory* of truth would be to follow the way of analogy: theoretical statements are no different from other statements outside the realm of science; a sentence like "electrons have spin ½" does not require any special, philosophical analysis of how it is to be "literally" understood anymore than does the sentence "the picture on the wall is crooked." Both of these have something in common – a subject (a picture, an electron) has a certain property (being crooked, having spin ½). Failing that, if one feels the need to give a special explanation not just of theoretical truths but of *truths in general*, the correspondence theory is a good candidate: it provides a very robust account of how our words hook up with the world. We can presume that a literal interpretation of a scientific theory precludes reading it as a useful fiction or conceptual metaphor (as would the instrumentalist), or as a sort of elaborate code whose "real" meaning asserts something other than what it appears to say at face value (as would a reductive empiricist). But blocking these two anti-realist

positions is not yet to say why we *really need* a correspondence theory to explicate what theories "literally" mean.

There is another, more contemporary anti-realist interpretation of the meaning of theoretical language which, among other things, recommends a particular understanding of the truth predicate. This is "internal realism," first defended by Putnam (1981). In short, Putnam famously turned on his earlier metaphysical realist position because found it very mysterious that our concepts "hook up" with the world in such a tidy way that seems to carve nature "at the joints." His explanation of our apparent referential success is Kantian; our concepts are not fashioned *by us* to map the contours of a preexisting external world. Rather, the structure of the world is a function of (dependent upon) our system of concepts. He uses the metaphor of a layer cake; as it turns out, the cuts in the cake line up so that the cuts on the top level are exactly above the cuts on the lower level. Such an exact match would (allegedly) be a mystery if the bottom was already carved, and we tried to make the cuts on the top layer line up with the ones on bottom. Instead, the cut goes from top to bottom. I.e., the cuts on the bottom of the cake, the world's structure, are a function of the cuts on the top, our concepts. Internal realism is obviously incompatible with a correspondence account of truth where the direction of fit between truth-maker and truth-bearer is world-to-mind rather than the other way around, as the internal realist would have it. That theory requires correspondence to facts, and facts just are the ways in which reality is structured.

Internal realism cannot countenance mind-independent structure either – it is incompatible with **M**$_{CI}$; the world is conceptually dependent upon our existence. The way that the world is *structured* (the categories that exist) depends entirely upon us. So truth has to be something other than match with a *mind-independent* fact.

The internal realist's best hope for a truth surrogate is *ideal warranted assertibility.* A sentence's being true consists in its assertion being warranted or licensed by an ideal completed science. The limit of inquiry, the best theory we could possibly construct, replaces matters of fact in the truth-making relation. So the truth of a sentence is relativized to our (best possible) conceptual scheme – if the sentence is ideally warranted, then it is true, and that is all there is to truth. Internal realism thus embraces an *epistemic* theory of truth, defining it in terms of (possible) human knowledge. According to epistemic theories, it is metaphysically impossible that there be a gap between what we know in a best-case epistemic scenario and what really is the case. Contrast this with the correspondence theory, where the epistemic notion of evidence or warrant is divorced from truth. On the correspondence theory, there might be sentences or propositions made true (or false) by a particular fact, *even if* all the evidence we could possibly gather indicated otherwise. Again, we see how epistemic theories of truth provide a solution Putnam's original worry about how our concepts could possibly track the world's structure. Epistemic theories of truth go hand-in-hand with a rejection of metaphysical realism: it is instead our epistemic practices that are rock bottom

fundamental (for all we know). The world, and true statements about it, are derivative from these practices.[3] This dissolves any worry that our theories could in principle mismatch the world.

In short, the correspondence theory of truth would certainly rule out internal realism, and this is one strong reason for realists to embrace it. The correspondence theory quite obviously makes truth *evidence-transcendent*. Insofar as scientific realism involves **M**CI, the idea that there is a mind-independent world described by our theories,[4] scientific realists *must* allow for the possibility that our best theory of the world is false. Otherwise, our theory really would determine the world's structure, and not (as realists believe), the other way around.

There is a final way in which scientific realism has implications for a theory of truth. This stems from the realization that (**a**) any theory is likely to have some false content, because scientists are continually improving their theories and in many cases *know* that their theories are inadequate and incomplete in certain respects, and (**b**) that some of our very best theories are mutually inconsistent – quantum mechanics and general relativity are the prime example – and therefore our total theoretical "world picture" is necessarily incomplete. Realists *have* to admit that, strictly speaking, all of our

---

[3] Internal realists do not go as far as to say that there would literally be no world if there were no world-for-us; that if we were to vanish, so would the rest of the world. Instead, one has to think very much like a Kantian: the only world we can ever know is the world-for-us that is fashioned entirely out of our concepts and perhaps implicit practices. The concept of a "world in itself" is only useful as a limiting concept of what we *cannot* know; it is if anything an amorphous "blob" upon which our concepts impose a structure.

[4] This is just another illustration that the connection between semantics and metaphysics is still quite tight even though they are strictly speaking distinct.

theories are false. They are instead said to be false but "approximately true," "verisimilar," or "truth-like." This was hinted at in the passage where Boyd says theories are "confirmable as *approximately true*." It is pretty clear that a scientific realist who wants to characterize his position in terms of truth also needs to accept the idea that theories are approximately true, "much more right than wrong," but still strictly speaking false.

This idea of approximate truth is intuitive enough, and seems to apply to ordinary cases where we need no background philosophical theory to elucidate what we mean by "approximation." An example: I am (roughly, approximately) 5' 8 1/2" tall. I have never measured my height on a more fine-grained scale, but it is a near mathematical certainty that my height is not *exactly* 5' 8 1/2"; instead it is a little less or a little more than this exact value. Nevertheless, it seems "more correct" to say that my height is 5' 8 1/2" than it is to say that I measure 5' 7". The former statement is closer to the truth than the latter, in a precise numerical sense: my actual height's difference from the former value is less than it's difference from the latter one. And comparative judgments are possible; it is still "more correct" to say I am 5' 7" than it is to say I am 4' even. We can rank statements according to their verisimilitude, or "closeness" to the truth. Here is a clear sense in which sentences can be strictly false yet truth-like. When one realizes the full extent of vague predicates in our language, there is a real *need* for truth-likeness. It can be more or less correct to call someone tall, or bald, or pale. Truth-

likeness is clearly an important (but comparatively neglected) semantic notion. If we can provide a philosophical analysis of the concept, perhaps it can be extended to the (scientific) theoretical case. So finally we have:

> [**S**VT (*t*)] *Verisimilitude Thesis:* Scientific theories *t* that are deserving of realist commitment are not true; strictly speaking, they are false. However, they are very close to the truth; they have a high degree of verisimilitude.

However, attempts to precisify this intuitive notion and to provide a logic of truth-likeness have had at best mixed results (see Psillos 264-5). At any rate, the Verisimilitude Thesis remains ambiguous, indeed mysterious, without a robust analysis of truth-likeness. It seems obvious to me that the intuitive account of comparative magnitudes (e.g., my height example, or counting hairs on people's heads) won't work in the case of theory. Approximate theoretical truth is strongly disanalogous to the ordinary cases of verisimilitude that involve estimation and vagueness. This is seen by looking at what it takes for one statement to be more truth-like than another. In the height example, we can move in the direction of increasing verisimilitude by introducing a finer grained scale for measuring height. We "tweak" the value of 5' 8 1/2" in one direction towards the actual value. For vague predicates where *some* quantity can be found, counting or other quantitative measurement (e.g., frequencies of sound,

shades of color) will do. In the absence of quantitative guides, say, in comparative judgments of who is more beautiful, or what food tastes better, we will need to adopt some form of arbitrary but standardized conventions for arriving as intersubjective judgments as possible (in practice, beauty and taste are just far too subjective without a background of shared preferences).

The revision of theoretical statements has almost nothing in common with these "everyday" judgments of comparative verisimilitude. The switch from Newtonian physics to special relativity required introducing brand new theoretical concepts (space-time, rest mass, etc.) and law-statements (the speed of light is constant with respect to all inertial reference frames). In this change we see nothing like the "tweaking" of some numerical value to fit the facts better. We *do* see a sort of precisification of vague concepts; the original term "mass" was ambiguous between designating relativistic mass and designating rest mass, because this distinction was never made before special relativity. As Field (1970) says, "mass" *partially referred* to both. However, this linguistic change is not like the example of more or less true bald-ascriptions. In that case, (**a**) we know all along that the term is vague (unlike the Newtonian's "mass") and (**b**) the vagueness of the predicate "bald" consists in its applying more strongly to some cases than others, whereas the terms "rest mass" and "relativistic mass" are not vague in this way. All this serves to demonstrate that approximate theoretical truth is importantly

dissimilar from the "intuitive" cases where we don't need a *theory* of verisimilitude to make reliable comparative judgments of approximate truth.

Proponents of **S**VT probably mean by "Newtonian physics is truth-like (compared to Special Relativity)" something like, "Newtonian physics is a special case of Special Relativity in the limiting case where $c \rightarrow \infty$ (where relative velocities are small compared to the speed of light in a vacuum), and is therefore a reliable description of how the world works in those special cases." But then they will need to say what it is for one theory to be a special case of another. The "correspondence principle" made famous by Niels Bohr requires that the laws of an old theory are mathematically derivable from the laws of its replacement theory as some value tends towards some limit (typically zero or infinity). Newtonian physics satisfies the constraint from two different directions: it is recoverable from quantum mechanics if one lets Planck's constant tend toward zero, or from special relativity if one lets the speed of light tend towards infinity. In each case, however, the way in which an earlier theory can be considered approximately true changes. From the perspective of special relativity, Newtonian mechanics can only be considered approximately true because the speed of light is so great in comparison to most terrestrial speeds that it can be safely ignored. From the perspective of quantum mechanics, Newtonian mechanics is approximately true because the discrete quanta of energy are so small that energy can be treated as a continuous quantity, as it is in the latter theory. Although the mathematical notion of a limit is common to both, the

scientific details change. This is recognized by Psillos (276-79), who offers a sort of *contextualist* theory of truth-likeness.[5] A theory is truth-like only when it describes the actual world under what Psillos calls *specific conditions of approximation*, and the details of these conditions are to be filled in by either the theory itself or a background theory. On this view, judgments of truth-likeness are always theory-laden. We can only explain what it means for Newtonian mechanics to be false but truth-like by adverting to more theory (specifically, the superseding theory – quantum mechanics or Special Relativity) that explains *why it is* that the theory worked so well before in such a way that the theory *was not radically "off-track."* In the case of Newton's three laws of motion and both Special Relativity and quantum mechanics, this is relatively clear-cut, as there are very smooth limiting relations. But Newton's Law of Gravity is a notoriously tough case – although one can understand quite well from the perspective of General Relativity *why* the inverse-square law worked so well, there is no *gravitational force* in General Relativity, but instead inertial motions in curved space-time. Mathematically, one can see why Newton discovered an empirically adequate law. But given the background theory of General Relativity, it appears that his ontological posit of a force was just plain wrong and rather unlucky – unlike the case of his three laws of motion, which *do* need to

[5] Psillos calls his theory "intuitive" but I think "contextualist" is a much better descriptor. The essential idea is that truth-likeness cannot be spelled out in a formal theory, without knowing the content-specific details of two theories we are comparing for truth-likeness, or the theory-world relation in the non-comparative case of a theory that *almost* gets it right, but not quite. To understand how a theory is truth-like is to understand exactly where the theory goes "wrong" and why, and also, I would add, to have some grasp of the possible solutions to improve or change the theory.

be *modified* in the Special relativistic context (i.e., high relative velocities) and quantum-mechanical (i.e., position-momentum uncertainty), but not outright replaced, as they smoothly transition into the familiar laws within the classical domain.

Even though it has been mentioned frequently, I have not yet said much up to this point about *reference*: how it is that words come to be paired with individual objects, kinds, events, properties, etc., and what it is that makes a given word-thing assignment either *correct* or *incorrect*. More concretely, we want to know: (**a**) the etiological question; *how* did the *name* 'Morah' get "paired up" (so to speak) with a specific person, my daughter Morah, or how did the kind-term 'infant' get paired up with the class of infants, or how did the predicate 'cute' get paired up with either the class of cute things (or the universal of cuteness)? (**b**) The normative question; what *standards* or *norms of use* are in place, such that one can be *mistaken* about who Morah is or what cuteness is? Many scientific realists are happy to leave these questions alone, and that is just fine: this work is probably better left to those engaged in the philosophy of language, *given that* scientific language is supposed to be just as "literal" as ordinary discourse.

However, some scientific realists have offered their own interpretive glosses on the reference relation, as it appears in the correspondence-truth schema. Realists are quite fond of the "new" causal theory of reference of Kripke (1980), which displaced the "old" descriptive theory (broadly, the referent of a name is fixed by a description or some set of descriptions of the thing, of which the speaker is aware). The causal theory

instead says that terms get fixed by an initial dubbing or baptism (e.g., when a parent names a new child), and then that term's referential use spreads from speaker to speaker over time, with each new speaker "borrowing" the reference from the original baptizer in the past. (This explains how, at this current moment, I *know* that 'Anaximander' and 'Anaximenes' refer to two distinct people, even though *I* am predictably unreliable and confuse them all the time and could not provide a description that distinguishes one from the other).

Scientific realists have been especially attracted to causal theories of reference, as it is believed that they will help to secure some form of principled *continuity* of the subject matter and descriptive content during the process of theory change, perhaps even in rather dramatic cases where our understanding of the world is radically altered. This idea of continuity properly belongs to what I am calling a "developmental" thesis, as it concerns facts about the nature of scientific change over time. But the causal theory of reference obviously provides one very strong pillar of support for continuity across theory change: *if* you are a correspondence-truth realist, then, then if you "fill in" $\mathbf{S}_{CoC}$ and $\mathbf{S}_{CaC}$ with a *mechanism* for reference-fixation (a term's initial baptism and causal history), then *beliefs* (or theoretical descriptions) concerning things in the world can change, in fact they can change quite radically, without *reference* to that thing being severed: the word-world link remains stable, even though the word has now taken on a

whole new cluster of associations and descriptions that it never had before, and also lacks many of the associations and descriptions that it used to have.

Consider, for instance, changes in the *meanings* or the *descriptive content* of the term 'electron' over time. J. J. Thomson famously discovered that cathode rays were not really "rays" at all, but rather streams of charged "corpuscles" that eventually were named 'electrons.' Radioactivity was discovered around the same time, and it was soon figured out that "beta particles" had the same charge-to-mass ratio as the cathode particles, and so were instances of the same kind (this also suggested that electrons are quite possibly constituents of atoms). At this point in time, the term 'electron' referred to a new and exotic but still classically "well-behaved" point particle that obeyed the classical laws of motion and of electrodynamics. But with the advent of the new quantum theory, the same word 'electron' was now used to describe something very bizarre from a classical perspective – still particle-like, but now with dual wave-like (delocalized) characteristics, which does not trace out precise trajectories[6] through space, and even appears to "jump" from one location to the other without traversing the space in between. Closer to the present, the term 'electron' is used to refer to a particle within the standard model, still exhibiting bizarre quantum behavior and properties (i.e., it has

[6] For any particle to trace out a definite path over time in three-space, i.e., for it to "move through space" in accordance with the laws of motion, it has to successively occupy spatially contiguous points at temporally contiguous moments in time, *and* at each point it must have a well-defined momentum. But no (quantum) particle simultaneously possesses a precisely defined locations *and* momenta; the determinacy of one trades off the other. So electrons cannot have classical trajectories.

spin ½ in addition to the classically well-behaved properties of charge and mass). It is the lightest of the three other corresponding particles in its generation (the muon and the tau), and its electromagnetic influence on other charged particles is transmitted via the exchange of "virtual" photons. Furthermore, if quantum field theory is to be believed, electrons are not so much particles (with wave-like characteristics) so much as excitations in an underlying field.

Clearly, the descriptions associated with the term 'electron' – the "meaning" of the term, if you will, provided by theory – have changed so much that it appears a descriptivist about reference would have to say that Thomson's use of 'electron' could not refer to the same thing we refer to when *we* use the word 'electron.'[7] But we could instead say that the token utterances of the word 'electron' name the kind of particle that are the constituents of cathode rays and are identical with beta particles – a correct judgment, from the standpoint of today's researchers, it is just that we know so much more about what these things *are*, and they had many false beliefs and preconceptions about them. Then, reference and meaning become detached from each other in a way that allows for a change in our theory of that thing, ranging from mere addition to what we already know, but also allowing quite *radical* shifts in our beliefs *about* one and the same thing, so that it makes sense to say, e.g., "scientists used to believe electrons had classical trajectories and definite locations, but our best current theory now suggests that

[7] But for a dissenting view and a principled defense of the use of a descriptivist theory exclusively for theoretical terms, see Papineau (2010).

they really are quantized excitations of the underlying electromagnetic field." One virtue of the causal theory over a crude descriptivist account is that it seems to make sense of both false belief and changes of mind in a way that the crude descriptivist can't, as we have already seen – furthermore, as we will see below, the nature of theoretical change is another key issue of concern to scientific realists, and since a causal theory (versus a crude descriptivist theory) would make possible changes in wrong theoretical beliefs, while hopefully retaining our ontological commitments (and some important correct beliefs). While some scientific realists do not explicitly endorse any particular theory of reference (and this may just be because the causal theory is the default "received view" in the philosophy of language), others – see, for instance, Hardin and Rosenberg (1982), or Psillos (1999)[8] – make it a key component in addition to a correspondence understanding of the truth predicate. Thus, it is worth including the following thesis:

> [**S**$_{CR}$ (*t*)] *The Causal Referential Kind-Realism Thesis:* Theoretical kind-terms of a certain theory *t*, like 'acid,' 'up-quark,' 'gene,' and so forth, refer to entities in virtue of a scientist or group of scientists coming into contact with a paradigmatic sample of the kind either directly or indirectly, through the kind's

[8] To be fair to Psillos, his is a "hybrid" causal-descriptivist account that lets reference be fixed *initially* by what he calls a "core causal description," a description of the causal role of the postulated entity plays in a network of causal relations. *If* the description is satisfied by a unique entity that plays that causal role, then this is the referent of the term. Further uses of the term inherit their reference from the *initial description,* which then gets passed down from speaker to speaker *per* the causal theory. I mention Psillos because he relies on the stability or "preservation" of reference on a causal account to help along his defense of scientific realism.

characteristic effects, and by the scientist(s) making the decision that the term will henceforth refer to all and only all things that are members of that kind.

I have emphasized a few key points of the theory that are of special interest to scientific realists. The realist is of course most interested in *kind-terms* for *types* of objects, properties, relations, and events, as theories are always generalizations about the world. The typical mechanism for baptizing an entire kind is not by simple ostention (as when I point to, pick up, or otherwise gesture towards my dog and say 'his name is Beau'), but rather by coming into causal contact with a "paradigmatic sample" of the kind in question and dubbing all things of this general kind to fall under the extension of the term. There is an explicit metaphysical assumption being made here, namely **M**$_{NK}$ – natural kind realism. If the kind really already exists "out there," then there is no need to make any further decisions about the extension of the term *beyond* getting a genuine, relatively pure sample, because the "unifying property" – the *essence* or *kind-constitutive property (-ies)* – will serve as the identity criteria for kind-membership and hence being within the extension of the term. There is of course a very tight *prima facie* conceptual link here between one's metaphysics and semantics, as is generally the case with all of these various dimensions.

Specifically, I do not think one could endorse a causal theory of reference for *kind-terms* without believing that there are essences that fix a kind's extension,[9] due to the open-ended nature of kind-concept application; e.g., we might learn new things about what electrons are like, which means that our initial grasp of the concept 'electron' – which I would identify with our ability to *correctly identify* electrons – is *necessarily partial* when we first identify the putative kind (we never start with a complete theory of the entity under investigation, and otherwise there would be nothing to learn, and no reason to do science). But there can be no correctness conditions of application in place if there is no essence or set of kind-constitutive properties (on the metaphysical side of the coin) to fix the extensional identity criteria of the kind-term (on the semantic side of the coin); thus, there are no facts of the matter about what my utterances of the word 'electron' refer to – the word then just denotes nothing, just like the antiquated kind-terms 'witch,' 'miasma,' and 'quintessence.'

Therefore at least some *conceptually independent* boundaries existing objectively in nature *have* to exist: otherwise, the causal theory of reference will have nothing to appeal

---

[9] Notice I have nothing to say on the matter of reference to individuals, that is, *names*. Kripke of course famously argues that individual things (especially people) have essences as well; the essence of a specific, particular thing are certain facts about the way it came into being, and these facts depend on the specific type of thing that it is – for instance, most people very strongly believe (upon philosophical reflection, at least) that a given person would not be self-identical if they had different biological parents, or even if they had failed to originate from the same meeting of sperm and egg (a person's *genes* seem to constitute their particular essence). The essences of material objects, on the other hand (tables, boats, pencils, etc.), are the original matter used to make the object, combined with the particular way in which this original matter was arranged into that very object. For example, my car is not identical to anything that is completely *qualitatively* identical to it (dents, scratches, and all), because my car is essentially composed of particular, individual parts arranged in a certain way; consequently anything made of *different* parts, even if arranged in exactly the same way as my car, deliberately dented and scratched, and switched in the middle of the night, it still wouldn't be *my* car.

to serve as kind-constitutive properties/identity criteria for the kind: for there are no unifying properties, no essences. Thus any grouping of entities under a common term has to be conventional – certain sortal concepts may be *useful* – recall the term 'fragile,' which names objects that have a common superficial property but no underlying unique reason *why* this is so; even still, it is much more useful to know which objects are fragile than it is to know which objects are grue. Conventional kind-terms are only more or less *useful* for organizing the world, and are essentially relative to *our* purposes; they are not meaningfully correct or incorrect, since that would require objectively unified structures and real distinctions in nature. The upshot is that this theory of reference requires a rather robust metaphysical scheme, i.e. the existence of natural kinds and (to some extent) a "ready-made" world. Thus, it is impossible to avoid the conclusion that $\mathbf{S}_{\mathrm{CR}} \supset \mathbf{M}_{\mathrm{NK}}$.

There are two further nuances here. The first is the notion of "indirect" causal contact. It is one thing to say (as Putnam famously did) that the extension of the term 'water' is just fixed by whatever micro-structural property makes all samples of water similar – because we can see and interact with water (it should be noted that 'water' is most certainly *not* a theoretical term, although 'H_2O' is). Without revisiting the strict separation between observability and theoreticity, it should be clear that we are going to come into contact with many theoretical kinds without directly experiencing them. In this case, we dub a kind term through contact with the effects that it manifests – such as

when we identify a new subatomic particle in a particle accelerator, or even when we postulate a kind of phenomenon as responsible for many different kinds of effects acting through the same principle – as Newton did with gravity.

Second, Kripke and Putnam often speak of the "discoveries" of essences – of finding out what the *real* nature of water, gold, tigers, etc. is. When scientists identify a new kind – say, a new type of subatomic particle – they already have the working background theory to fill in the "unifying property" that plays the essence role, so it is known ahead of time what "glues" the kinds together (so to speak). But this is not always the case, and that is what Kripke and Putnam were onto. Some familiar objects, properties, and other types of natural phenomena might have these types of unifying properties (if one believes that there are natural kinds and kind-essences), but we may not yet have made the right discoveries or possess the requisite scientific concepts (e.g., an evolutionary lineage, supposing for argument that species are kinds with historical essences). The basic idea is that we humans have "essentialist" intuitions: we will group things based on their surface similarities with the belief that there is something that is responsible for these similarities even if we don't know what this is, and part of the business of science is to uncover these kind-constitutive essences.

Finally, we need to examine how the global-local distinction bears on all of these semantic theses. It is clear that the Conditional form of Correspondence-Truth Realism is global in the sense that the realist is offering a full-blown *philosophical theory* of what

makes any statement true, including those occurring within the context of science. There is just one theory of truth for all types of sentences: when the realist wants a theory to be understood "literally," at the very least, she is pleading for no double-standards in semantics, e.g., construing theoretical language as non-truth-apt and "ordinary" language as truth-apt.

It does not make sense in the absence of any strong reasons to offer a disunified account of the truth predicate for different domains (or sub-domains) of discourse. If, as many realists boldly assert, the correspondence theory is both true *and* needed to articulate or defend scientific realism, then this is the view we ought to adopt, for both non-scientific and scientific discourse. This is why I said it is the theory of truth in place "by default" when explaining *Conditional* Correspondence Truth Realism. However, I will leave the door open for localism, without requiring that correspondence being the "only game in town" for the simple reason that scientific discovery might uncover a theory for which there are good reasons to adopt a different truth-semantics.

I think the best case for a "localized" understanding of the truth-predicate based on distinctly theoretical considerations might come from a theory like the "Copenhagen" interpretation of quantum mechanics, the central feature of which is the denial that Schrodinger's "wave-function" is a physically real process; appearances to the contrary, it does not describe the state of some independently existing thing (like a Bohmian "pilot wave"), or process evolving over time. Instead, it represents the *state of our knowledge*

about the physical system being described. The probability of an event occurring *when we interact with (measure) the system* (such as a particle being found at a particular location) is given by the square of the amplitude of the value of the wave function associated with that event. The relevant sense of probability here is *subjective*; it represents the degree of belief we should assign to the proposition that the event occurs.

According to Bohr and Heisenberg, a subjective interpretation of the wave function (and so-called "state reduction" of the wave packet) is forced upon us an important fact: that we have no choice but to measure, describe, and think about the physical world by employing the *classical* framework of things existing at determinate positions in space and as capable of moving through this space with a definite trajectory. But neither application fully describes the nature of the micro-quantum world: Bohr's complementarity principle has it that photons, electrons, and so forth are paradoxically *both* particles *and* waves in the sense that each behavior appears to trade-off one another: the more localized the wave function is (a spike in amplitude near a spatial point), the more we know about something's position and the less we know about its momentum, and the less localized it is, the more we know about its momentum and the less we know about its position. As things have a dual nature as *both* particle and wave that we cannot fully grasp, we are confined to describing the object as *either* particle *or* wave – but again, only when we *look for it*.

On this interpretation of quantum mechanics, the theory does not attempt to describe the world as it really is in itself. In fact, Bohr and Heisenberg thought that we *cannot* do this, and this cognitive limit might properly be considered part of the Copenhagen interpretation. Since we can't fully understand the quantum world, the theory is essentially a *description* of the world insofar as can we know it. Its fundamental laws (the wave equation and the uncertainty relations) get interpreted as describing our *knowledge* of the system, not the system as it is *in itself*. Thus Bohr famously said, "it is wrong to think the task of physics is to find out how nature is. Physics concerns what we can say about nature."[10] This way of looking at quantum mechanics is inherently *constructivist.* The aim of quantum mechanics (on Bohr's and Heisenberg's view) is not *merely* to make true predictions, as, say, with instrumentalism. It is instead self-consciously cognizant of the "fact" (if it is a fact) that our knowledge of the quantum domain is not some pure, objective description of the world as it really is. It is constructivist in the sense that certain quasi-Kantian *a priori* forms of intuition make possible (and at the same time, delimit) the ways we can form a "picture," or *intelligibly represent* objects in space-time. Classical concepts must structure our experience before we can understand the world, and unfortunately for us, we can't apply these concepts except in a partial manner beyond a certain limit – very small scales of length and speed, or mass/energy and time (but where to draw this line is *notoriously* vague, and the

[10] Quoted in Petersen (1963).

source of many interpretive problems). And because quantum phenomena can only be partially described using one or the other "complementary" concepts, our human forms of understanding are inextricable from the (Copenhagen interpretation of) the theory.

The Copenhagen understanding of quantum mechanics, if true (and I am not suggesting that it is), would have tremendous metaphysical and epistemic import. But just as much, it would have implications for the *meaning* and *interpretation* of scientific language as well, as evidenced in Bohr's insistence that experimental situations provide the preconditions for intelligibility, and more overtly revealed in his comment on the proper task of physics pertaining to *what we can **say** about nature*. That quote seems to suggest some kind of non-correspondence reading: in fact, it seems to suggest a sort of reading that a person who embraces an epistemic theory of truth would be happy with: (at least quantum) physics is "about" our own knowledge, i.e., what we are licensed to "say" or conclude about nature but only as it is seen "from within" our inescapable theoretical constructs. I do not know if Bohr or Heisenberg would accept this semantic theory, and speculating about their personal views is fruitless, but I *do* think that an epistemic theory of truth, where truth is coherence with our best theoretical system, seems to accord very well with certain aspects of the Copenhagen interpretation. These are the more anti-realist elements: the subjective interpretation of the wave function as simply encoding our knowledge in probabilistic terms (its physical *unreality*) and the reticence to say anything about what makes something a measurement (i.e., what

"causes" state reduction, even if this is only a change in our knowledge of the system), or where the classical/quantum boundary lies. On the other hand, such a theory may not be appropriate for *every aspect of the theory*, e.g., Bohr's apparent realism about the *ontology* of the theory, such as atoms, electrons, even photons.

Suppose for the sake of hypothesis that no progress was forthcoming on the measurement problem and further developments in developing a theory of quantum gravity do nothing to fix the problem that we must apply classical concepts in a piecemeal fashion to describe a non-classical world. This is close to, but not quite, the view of the internal realist; Bohr did not disavow an external world, nor did he think that we somehow *actively* structure it with our concepts, but he did seem to think that our *knowledge* of the world needs to be mediated by certain *a priori* forms of intuition which unfortunately just do not apply on the micro-scale. Still, consonant with the view that the task of physics is to describe "what we can say about nature," in this very specific circumstance one might want to give up the notion that truth is correspondence to a fact and instead say that whenever a proposition about some quantum-mechanical state of affairs is uttered, it is true just in case it is licensed by the axioms of the theory (along with the measurement postulate and Born's rule for subjective probabilities). This would be one way to be selectively *anti-realist* about (perhaps selective parts of) quantum mechanics: to simultaneously give up the metaphysical view that the quantum mechanical world really is the way the theory tells us it is, apart from the human

knower,[11] epistemological view that physics gives us *knowledge* of what the world is really like (because we know that we have a distorted picture through the lens of our classical concepts), and also the semantic idea that truth is correspondence to a worldly fact (rather than agreement with a theory). By adopting a local anti-realist understanding of the truth predicate for quantum mechanics, the would-be correspondence-truth realist is essentially "cordoning off" a part of nature as unknowable at least as it is in-itself, and thus in need of a special *dispensation* from ordinary semantic norms that *are* in play where there scientific realist wants to make her stand, on *other* theories that should be understood "literally." But again, these dispensations can only come from very special, topic-specific theoretical considerations. Otherwise, only one philosophical theory of truth is in play when we analyze statements, both theoretical and non-scientific – and if **S**$_{\text{CoC}}$ is right, that theory is the correspondence theory.

Now, of course, the Categorical form of the Correspondence-Truth Realism Thesis is of course very local in the not-very-exciting sense that (probably) some theories are false, and some theories are true, and it will depend on the details of which is which. It will *still* be the case that for each theory that *is* true, generally, a realist will want to say that it is true in a correspondence sense. The same goes, *mutatis mutandis*, for **S**$_{\text{VT}}$, the

[11] Although this may not be entirely correct: if the theory is taken to be the bare formalism of the theory *plus* the measurement postulate, then yes, one might think that the "job" of the theory, as it were, is to describe nature in itself. If we add the Copenhagen gloss and take *that interpretation* to tell us the full content of the theory, then the theory *actually discloses its own instrumental nature* and the ineluctability of the role of the observer.

Verisimilitude Thesis. Different theories will have differing degrees of truth-likeness, and especially given the contextual nature of verisimilitude that I have offered here, this thesis is even *more* local as a matter of necessity. Not only will specific theories differ in whether they are truth-like or not, but they may also differ in their degree of truth-likeness, and most importantly of all, they will differ in the *context-specific details* of what *makes* a certain theory truth-like.

Last of all is the question of how global or local our theory of reference should be. I think that there is some sense in saying that perhaps a causal theory is inappropriate for at least *some* scientific disciplines *if* one is not inclined to believe that the kinds found within that science are *natural*, are organized by objectively existing differences in the natures of things, and not human conventions. This attitude may vary across scientific disciplines.

For instance, some philosophers of biology take an *eliminativist* (or error-theory) attitude towards the species concepts[12] I mentioned in the context of **M**NK. Our ordinary intuitions about essentialism fail us; there is no unifying property that carves out biological "kinds." There are only our conventional ways of grouping organisms into nominal species, which may be more or less principled but not ultimately *correct* according to our practical scientific needs. There are, of course, deeper reasons for the surface similarities of organisms – namely, evolutionary forces such as selective pressure

[12] See Ereshefky (1998).

and genetic drift, but one will not find anything like a kind-constitutive property or set of properties (not even a principled *historical* essence). One might say all this and still take (for example) a robustly epistemological realist stance towards evolutionary biology: we can *know*, for instance, that all life on Earth has descended from a common ancestor, and we can even reconstruct (some of) the branching "tree" of life from this common ancestor all the way up to the present through analysis of the genetic structures of living organisms. It is just that biology has no natural kinds to speak of or generalize over. For any two organisms, we may be able to tell how near or far they are to each other on the tree of life, but there will be no *sharp* fact of the matter concerning whether they are of the same general kind – there are no kinds at all, only individuals.

Whether one personally agrees with this attitude or not, suppose that it is true (as some in fact believe). If one can grant that, then one can understand how the causal theory of reference might fail due to theory-specific reasons, even if a causal theory of reference is appropriate most everywhere else, say, in chemistry and physics where there are very clear kind-constitutive properties. It is just that in our past "folk biology," we misclassified populations of phenotypically similar animals as belonging to a unified kind whose boundaries are sharply delimited by the common possession of some hidden essence. Knowledge of common descent erases the possibility of any essential traits, thus making our innate assumption that there is a real natural kind corresponding to each species term just *wrong*, and therefore species-talk can only serve a conventional

purpose. Whenever a biologist groups individuals into species, it is for principled but ultimately pragmatic reasons. Perhaps then one can "borrow" this species term or concept and refer in virtue of whoever first devised the concept, but the main thrust of the Referential Kind Realism Thesis would lost most of its thrust. Kinds are not delimited by real boundaries set by nature, and nothing prevents another biologist from adopting another classification scheme that is better relative to a different purpose. But again, consistent with this, one could maintain the Referential Kind Realism Thesis for other sciences – specific facts about the very nature of biology make it so that, in this particular instance, we should expunge science of talk of natural kinds. If this position is consistent and defensible (as I believe it to be; whether it is true or not is a separate but irrelevant question), then it is also consistent and defensible to adopt a localist attitude towards **S**CR.

5. The Priority of Epistemology

Having surveyed the various proposals for definitive statements of scientific realism, the "descriptive task" of this project is complete. Now I turn my attention to the "normative task" of determining which of these proposals are worth taking seriously, and which are orthogonal to the central issues. Here I am going to make a *prima facie* case for epistemology as the "core" of scientific realism.

The goal of this project is not to convince anyone that we *should accept* scientific realism in any of the above senses. Before one can argue for a philosophical position, one needs to get clear on what one is arguing for – and scientific realism is less a determinate philosophical proposition than it is a loose association of various positions that are often confused with one another, owing largely to historically contingent reasons. The conglomeration of these different attitudes is, I believe, a historically contingent *reaction* that can be traced back to the downfall of positivism and its reductive empiricist and/or instrumentalist attitude towards scientific theories, which left an interpretive vacuum.

As I see things, the revived interest in robust semantic realism was, for instance, in a substitute for the verificationist theory of meaning and truth. In the philosophy of science especially, programs to reduce theoretical language to its pure empirical content, or to dispense with talk of meaning and truth altogether by treating theoretical statements as an uninterpreted calculi failed miserably. At the same time, Thomas Kuhn's vastly influential *Structure of Scientific Revolutions* (1970) cast doubt upon

triumphal ideas of scientific progress and the rationality of theory change; Feyerabend and other post-modernist thinkers took up Kuhn's talk of paradigm shifts, and spun his rhetoric of revolution and social negotiation into strong forms of social constructivism. Thus we see scientific realists reacting by insisting that theory change is a rational process, continuous with and building upon what came before. We also see robust forms of metaphysical realism intended to ward off the more bizarre consequences of strong constructivism. To those who opposed these new developments, the new causal theory of reference was a godsend. With reference to natural kinds secured, independent of false beliefs about them, continuity and reference-stability are restored. Notice just how deeply interwoven these developments all are, but also consider how the specific way in which they interact carves out an internally consistent, perhaps quite plausible worldview – depending on your antecedent beliefs, of course. This worldview, this specific, very interwoven association of semantic, historical, metaphysical, and epistemological beliefs, I contend, is largely a reaction to these emerging threats to the *rationality* of science, which was something even the positivists appreciated. Scientific realism came to be associated with so many different viewpoints that it seems at one and the same time stifling (demanding very rigid commitments in one's ontology/metaphysics, epistemology, view of history…), or, paradoxically, also *indeterminate* in content, to the extent that one is not careful to appreciate and explicitly accept the reasons for all of these deep interconnections between very different

philosophical disciplines. This contributes a great deal to the sense of "ennui" hanging over the whole discussion, as Callender and Magnus (2004) so aptly describe.[1]

Instead of giving an argument for scientific realism, my aim is to give a statement of it that pins it down as a relatively determinate thesis, one that makes the most sense of the arguments that have been offered for it (based on a principle of charity), but also explains why the thesis has all of these other associations. The way forward is to treat "scientific realism" as *primarily* an epistemological thesis, as the idea that we can and in fact *do* know some of the things high-level theories have to tell us about the world – not necessarily *everything*, but some things. Some minimal views in other disciplines will, *of course*, follow from this. One cannot totally isolate one's epistemology from one's metaphysics, since at the very least the former will necessarily bear on what one is warranted in believing about the nature of the world regarding the latter (to use a very salient example). But my contention is that the connections are not, or *need not*, be as tightly interwoven as they are almost always assumed to be. This is why I characterize my quest for the "real" realism, if you will, as a search for a *core commitment* – a sort of *credo* that is minimally necessary and sufficient for belief in scientific realism.

---

[1] Callender and Magnus diagnose the cause of this "realist ennui" as residing within *how we argue for* realism. They argue that both the no-miracles and pessimistic induction – that global arguments (their term is "wholesale," as opposed to "retail") commit a seductive but familiar error of statistical reasoning of ignoring base-rates of successful theories and of *true* theories out of all the theories that have been formulated. I fully accept their conclusions and I am willing to grant that a great deal of the sense of stalemate is due to these conflicting intuitions that are in the end full of sound and fury, signifying nothing. However, it is my goal here to identify perhaps an even more troubling fact – that scientific realism is not (and has never been) very determinate in its content, representing a whole constellation or association of views more than a single, unified coherent philosophical thesis. This also contributes to the sense of "ennui" and furthermore explains why we see such profound differences in approaches different philosophers take towards realism.

Intuitively, I believe scientific realism as a pure epistemological thesis makes a great deal of sense. After all, scientific anti-realism is, I think, best framed as essentially *selectively skeptical* stances on the possible and actual extent of human knowledge. There is no *a priori* reason to see the project of scientific inquiry as somehow set apart from more mundane projects of fact-finding (e.g., history, or detective work) in either the methods it uses or in the types of questions it tries to answer. True, its methods are distinctive in that they more refined, cautious, and self-critical, but this should serve to recommend its deliverances rather than impugn them, and yes, science does probe deeper than ordinary inquiry into the very fundamental and very general causes of what we observe, but without an argument to the contrary, this is a difference of degree rather than kind. Any argument for scientific anti-realism is, on my view, best seen as supplying that missing argument for drawing a distinction between scientific and "ordinary" forms of inquiry, and any positive scientific anti-realist worldview (whether empiricist, constructivist, or instrumentalist) is an attempt to explain exactly how the scientific knowledge we *seem* to have is not really knowledge, after all.

An analogy will help to bring out this point: the Cartesian skeptic is an external world anti-realist. She finds arguments for calling into question *all* of our knowledge, *all at once* – perhaps you are dreaming, or perhaps you are just a disembodied brain, floating in a vat, with a mad scientist electrically inducing simulated experience that you mistakenly take to be real. Such a skeptic always starts from *within* a default position of

our t*aking ourselves to know* things about the world (at the very least: that we sometimes dream things that aren't real, or that our experience of the world is mediated by neuronal activity) in order to *undermine* that putative knowledge. Before we ever engage in philosophy, we find ourselves in a default epistemological position of believing ourselves to know many things, and the skeptic attempts to move us *away from* that starting point. Similarly, I am proposing that science is just a proper part of our more encompassing knowledge of the world – perhaps distinctive in its precision, depth, instrumental value, etc. – but not somehow transcending the more ordinary knowledge-gathering activities that we engage in. *This* is our epistemological starting point: the only salient differences between me determining that a loose hose in my radiator was probably the cause of my car overheating, and a scientist's measurement of a photon's spin orientation are in the types of instruments we use and procedures we follow, along with the number and depth of the background assumptions we have, *but not* in the inference forms we use or our standards for evaluating the rationality of our beliefs, given the evidence that is relevant to the hypotheses we form. That is, in the absence of good reasons, the Principle of Epistemic Parity (**EPP**) holds. A *scientific* (as opposed to *philosophical*) anti-realist attempts to move us away from this starting point, by offering us reasons to think that we really don't have at least some of the scientific knowledge we took ourselves to have, because there is something distinctive about science or some specific scientific field that makes its conclusions untrustworthy. The essence of

scientific anti-realism is the denial of **E**$_{PP}$, whether in global form or in more local contexts (selective skepticism about certain scientific fields of inquiry).

This is not to say that all the scientific realist has to do is to offer a refutation of the scientific anti-realist's reasons for selective skepticism, any more than defusing all of the arguments for philosophical skepticism would *automatically* provide us with conclusive reasons to believe that an external world exists, without further argument (but see below for more on this). Continuing the analogy, it is not enough for an external world realist to defeat the skeptic. That would just remove whatever doubts we may have about what we take ourselves to know. But the external world realist seeks something more, namely, secure foundations for affirming that what we believe ourselves to know, we really *do* know. This is a kind of *second-order* affirmation of our knowledge. The default epistemological position is not one of philosophical reflection – it is where we begin from when we consider both skeptical proofs and vindications of a real, external world that is the object of our knowledge (or it is where we "stay" when we are not persuaded by either argument – see below).

Now, similarly, scientific realism has got to be something more robust than just the denial of scientific anti-realism – because this too is a pre-philosophical attitude that begins from a place of relative philosophical complacency rather than second-order reflection about the limits and extent of our knowledge. What makes scientific realism distinctly *philosophical* over and above arguments offered by practicing scientists is that it

offers some *positive reasons* to think that not only is there no in-principle limitation on our scientific knowledge, but that we actually *do* have at least some of this knowledge – that is, I believe scientific realism is equivalent to **E**TK, the Theoretical Knowledge Thesis. *Considered as epistemological theses*, skepticism and external world realism are logical contraries, not contradictories. Obviously, speaking metaphysically from a "God's eye point of view," either we have some knowledge of external world or we don't, and in this sense epistemological anti-realism and realism are contradictories. But backing up one step to examine our *reasons* for adopting either of these two epistemological standpoints, there is a *tertium quid* – namely, the "epistemological starting point" of pre-philosophical reflection, where we haven't yet either considered or decided whether there are better reasons for accepting one or the other thesis about our knowledge. The same is true for scientific realism and anti-realism: they are pro- and con-attitudes we can adopt towards a special subset of our knowledge, but we cannot properly be said to properly hold either view before we have delved into epistemology – the distinctly philosophical project of inquiring into the reliability, limits, and extent our own knowledge.

On my analogy, the "epistemological starting point" corresponds to a *non-philosophical* view of our knowledge, both mundane and theoretical for the global skeptic, just theoretical knowledge for the selectively skeptical scientific anti-realist. But what does this distinction between taking up a philosophical versus a non-philosophical

epistemic position amount to? There are two different ways that I believe we can occupy such a non-philosophical position. The first and much more common way is simply (as I have already said) is to occupy our "natural" point of view, before we have engaged in serious reflection concerning whether we know what we think we know. This is, generally speaking, an *implicit* epistemological realism. We don't, generally speaking, engage in *distinctly philosophical* worries about whether the loose radiator hose *really was* the cause of my car overheating, or whether the electron *really does* have spin in the upwards direction. We do, of course, realize that there are sources of mistakes and error both in methodology, instrumentation, and in inferential strategies, and if we have sufficient reason to be concerned about these sources of error, we double- and triple-check these things, or we seek means of independent confirmation to corroborate our conclusions. If we are rational, we humbly recognize our fallibility and we take a more-or-less attitude towards different pieces of knowledge, although it is psychologically unrealistic to expect that these could be assigned to sharp or even fuzzy ranges of subjective probabilities (as the Bayesian would have it): I am more confident that my car is blue than I am about the loose radiator hose hypothesis; the scientist is more confident that photons exist than that this particular measurement is correct. But these are all what I have called "first-order" considerations: the ordinary standards of rationality, method, evidence, what counts as an adequate explanation, etc. But, given all of these qualifications, we *do* think that we know some things, perhaps even quite a lot. (That is

why I characterized skepticism as the worry about whether we *in fact* know "what we ordinarily take ourselves to know").

"Distinctly philosophical" skeptical worries transcend these ordinary processes of belief-formation and of ordinary doubt. The globally skeptical hypotheses undermine all rational inference at once: no amount of re-checking our methods, assumptions, etc. will help to escape the brain in a vat hypothesis if our experience on which the inference is founded is illusory, nor will corroborating our putative knowledge by seeking independent evidence – for any agreement we find could simply result from the fact that the mad scientist has constructed a very elaborate and self-consistent illusory experience. The skeptical hypotheses are explicitly designed to rule out any ordinary method or epistemological criterion we might appeal to.

The scientific anti-realist's arguments (from underdetermination, pessimistic inductions, the greater epistemic security of empirically adequate theories over "true" ones, epistemic relativism, the problem of unconceived alternatives, and so on), if correct, are *distinctly philosophical* worries stemming from something *different* about scientific theories such that the ordinary procedures for eliminating systematic errors that scientists follow won't do anything to alleviate them. To pick one example, historicist arguments for epistemic relativism start with the Kuhnian idea that scientific controversies cannot be settled by appeal to some theory-neutral canons of rationality and method, nor will "nature itself" ("whatever that is") settle them, because paradigms

mediate the *way we interpret the world*, all the way down[2], and once a scientist has committed to a paradigm it is very hard to escape its clutch. This is the phenomenon of *paradigm incommensurability*: there is no shared language or conceptual scheme between rival theorists, so any argument for choosing one theory over another must necessarily appeal to assumptions and standards that are *internal* to each theory, so any choice between the theories has to be determined by extra-scientific, perhaps *non-rational* factors.

Epistemic relativists[3] fill this gap with the non-rational factors of social negotiation, politics, rhetoric and persuasion – factors which Kuhn himself suggested with his metaphors of scientific change as political revolution and scientist's declaring allegiance to one side or another as choosing sides in such a revolution, based on prior commitments, the influence of one's peers, etc. *Prima facie*, one can't appeal to our internal *scientific* methods and standards already in place to dispel this worry, just like one can't *prima facie* appeal to the our *general* epistemological resources to dispel global external world skepticism. That is just what incommensurability does: it raises the possibility that contemporary scientists are "stuck" in a historically relativized Kantian *a priori* scheme of theoretical concepts that structure our experience – but don't in any way

---

[2] I will ignore the stronger and mysterious thesis that paradigms somehow "make" the world we live in, which some of Kuhn's more radical remarks seem to suggest.

[3] As opposed to more rationalist philosophers like Lakatos and Laudan, who recognize the problems Kuhn raises about the theory-ladenness of theory choice itself but rely on the comparative ability of "research traditions" at "solving puzzles" to reintroduce an external, *extra-scientific* standard for theory-choice, thus restoring rationality to scientific development.

enable us to objectively represent it. These schemes can't be vindicated "from the outside," so to speak, and can only change by "jumping" to another paradigm on the basis of non-rational factors. So any of our *scientific* methods for conducting experiments, evaluating data, etc. and norms of what counts as a convincing solution to a "puzzle" are internal to this scheme. There is no escaping this problem if one grants the epistemic relativist's argument.

The second way of occupying the epistemological starting point *is* to be aware of the philosophical arguments, pro and con, but to be unmoved by either of them – and thus to adopt a view that refuses to take sides, that remains at the level of first-order questions and eschews any philosophical examination of our knowledge (either global, or restricted to scientific theories). These correspond to *non-realist* views in philosophy. Epistemological non-realism is not quite the same as the epistemological starting point since one is *aware of* distinctly philosophical considerations for and against a meta-justification of our knowledge, but it is characterized either by a reluctance to engage in these questions in the first place, or being of the opinion that either types of arguments fail to persuade. Non-realism about our knowledge of the external world in general is quite hard to find, but perhaps the one example might be found in David Hume (1999). Skeptical worries about the reality of causation "in the objects" themselves, or the endurance of self-identity over time, or even a real world corresponding to the causes of our experience followed from Hume's general theory of ideas and his famous "fork,"

which divided all knowledge into contingent (synthetic) matters of fact and necessary but trivial (analytic) relations of ideas. Hume recognized just how radical this skepticism was, threatening the very core beliefs upon which all our knowledge is founded – for instance, even if Descartes' proofs of an external world failed, he might still fall back on the *cogito* proof that he is a indestructible, non-material *res cogitans* that is the enduring substrate of all his thoughts and sensations. But Hume's incisive skeptical arguments against the reality of the self undermine even any justification for this position.

At the same time, Hume recognized that as a matter of pragmatism, one couldn't lead a life of hyperbolic doubt, even though philosophy tends to show us that we have no secure foundations for our putative knowledge whatsoever. As much as one might be convinced by his skeptical arguments against causation, we still think (as a matter of "custom") that it is far better to take the stairs than to jump out the window due to our past experience of all objects, everywhere falling to the ground when unsupported – even though there are *no* inductive rules of inference that can lead us to believe that this is projectable into the future – it is mere "animal instinct." His is a "mitigated skepticism" that is self-critical of the very doubts it arrives at – even though he thinks these problems are insuperable, they should not in any way affect our ordinary practices due to our *non-rational*, *instinctual inability* to distrust our senses (or to suspend judgment on causal necessities, or deny that I am the "same" person that I was twenty years ago) that we cannot but help as soon as we stop philosophizing:

> Indulge your passion for science, says [Nature], but let your science be human, and such as may have a direct reference to action and society. I prohibit, and will severely punish, by the pensive melancholy which they introduce, by the endless uncertainty in which they involve you, and by the cold reception which your pretended discoveries shall meet with, when communicated. Be a philosopher, but amidst all your philosophy, be still a man (89-90).

So Hume can be seen as feeling the force of skeptical arguments at the level of abstract philosophical concerns, but also as *remaining unmoved* by them due our lower animal faculties that we could not ignore even if we wanted to, and that we forget about as soon as we set abstract reflection aside so that we may return to practical affairs. We don't quite end up with an implicit first-order *epistemological realism*, back to where we started, but for practical purposes, we return to a kind of uncritical, unreflective trust of our senses, of causation, and our sense of being the same person over time – out of practical necessity, not out of any "distinctly philosophical" grounding. That is why Hume (quite paradoxically) calls this a "sceptical solution of these doubts" – there's no way to escape these "endless uncertainties," but so what? We *can't* help having these beliefs – it was just wrong to seek a philosophical, ultimate justification for them instead of locating them in our animal instincts.

Analogously, non-realism in the philosophy of science is the same general avoidance of meta-level epistemological questions that go beyond our ordinary scientific standards of theory appraisal, or a sense of doxastic equipoise between the realist and

anti-realist intuitions that tug in different directions; skepticism about a possible resolution of the equally compelling arguments. This is often referred to as *philosophical quietism*, the view that philosophy just has nothing to contribute to our "epistemological starting point," that skepticism and epistemological realism are equally implausible, and that as a matter of pragmatism, we ought to keep using our ordinary epistemic practices as a matter of *pragmatism* (as Hume endorsed) but not *inflate* them into a general philosophical view about our knowledge of the world.

The most popular proponent of this view is Arthur Fine, who defends what he calls the "Natural Ontological Attitude."[4] As I have argued, Fine agrees that realists and anti-realists begin with a "core position" that both realists and anti-realists try to *illegitimately* transcend. The core position just consists of our bare first-order beliefs within the framework of science. He accepts the Principle of Epistemic Parity (thus rejecting anti-realism) and even that he trusts the scientist when she avers belief in "molecules, atoms, y/J particles, and who knows, maybe even quarks" (Fine, 1986). So it would appear that fine accepts **E**TK and maybe even **M**FT, the Factuality of Scientific Theories Thesis – both epistemological (trust in what scientists say) and metaphysical (that there really are such exotica) assertions. Fine even seems to endorse a kind of semantic realism, when he "take[s] it that we are to treat truth in the *usual referential way* … NOA … commits us, *via truth*, to the existence of properties, individuals, relations,

[4] Another recent proponent is (I believe) Simon Blackburn, who has expressed pessimism about the debate in general and a criticism of the "no miracles" argument, claiming that there is "no getting behind the explanation."

processes and so forth referred to by the scientific statements *we accept as true*" (Fine (130), emphases mine). This would seem to be a correspondence theory *prima facie*, but Fine insists that his is a "no theory" of truth, refusing to take a stance on what the truth predicate, since both realists and anti-realists add on to this predicate (I think this is best understood as a deflationary, minimalist account, but then again, that would be a *theory of truth* rather than a refusal to answer). What the anti-realist adds on, according to Fine, is a non-standard theory of truth amenable to her anti-realism, for instance, a relativist theory that identifies truth with what is licensed by the set of social practices (a Kuhnian paradigm) currently in play, or an epistemic theory of truth favored by the so-called 'internal realist'. The realist adds, notoriously, "a desk thumping, foot-stamping shout of 'Really!'" (129), as in, "HEY, listen up! There *really are* electrons, *really! I mean it!* [THUMP]." What this table-thumping amounts to has been a source of confusion among interpreters of Fine, and Alan Musgrave has interpreted NOA as a disguised form of realism, since Fine does seem to accept certain existence claims and the usual semantic framework employed by the (semantic) realist Musgrave (1989), and the "thump" might just signal Fine's distaste for an overzealousness on the part of the scientific realist.

But I think we can be more charitable to Fine. According to him, both the scientific realist and anti-realist appeal to extra-scientific standards of justification, going beyond the normal explanations scientists offer in the course of their ordinary practices. Both the pessimistic induction and no-miracles argument are illegitimate from this

standpoint. We can do no better to explain why two electrons obey the Pauli exclusion principle – cannot occupy the same quantum state – than to start from the fact that electrons have spin ½ in order to derive the result that their behavior is described by Fermi-Dirac statistics instead of the Bose-Einstein statistics of force-carrying particles. It is otiose to appeal to an extra-scientific standard, such as the idea that the theory that licenses this derivation is mature and successful, and therefore highly likely to be true in most of its detail. Similarly, an anti-realist add-on that this derivation is paradigm-relative, or only warrantedly assertible at the limit of inquiry, armed with a final, complete science, is unnecessary and adds *nothing* except unwarranted metaphysical or epistemic assumptions. He suggests we stay at the "core position," accepting neither argument, and appeal only to this first-order scientific explanation for why electrons obey Pauli exclusion.

I do not wish to dwell much on philosophically quietist views like Fine except to note their relevance to what I have called the "epistemological starting point," but to help the reader understand, we might make an analogy to Carnap's defense of an internal/external distinction advanced in his legendary "Empiricism, Semantics, and Ontology" (1956). As the reader may be well aware, Carnap famously endorsed the view that we can make assertions about the world within conceptual "frameworks" that are established by stipulating a system of norms about how language is to be used, what counts as a legitimate inference form, and so on. There are frameworks for our natural

belief in the existence of external objects (presumably including the theoretical entities of science), or universals (properties), or abstracta such as mathematical objects, sets, even propositions.

For Carnap, only questions *within* a framework are legitimate and make inquiry possible. Ordinary *assertoric* claims *within* a framework – e.g., "Beauregard is a dog," "2+2 = 5" are truth-apt (the former true, the latter false), and their truth is determinable only with reference to the norms that govern the use of the framework. But *existence questions* and questions of *truth outside the framework* – "does the number 2 *exist*?" or "are unicorns *real*?" and "is the sentence 'electrons exist' *true*?" are either trivial or illegitimate. The system of frameworks lays down the linguistic rules for the proper usage of terms and possible inferences. Here we arrive at the famous internal/external distinction. "Internally," agreeing to use these framework-constitutive norms, the questions are trivial. According to the framework of numbers, we agree to use the word '2' in accordance with certain rules (such that "2+2 = 5" is false according to these norms), and trivially, 2 "exists," since to exist in an *internal* sense is to be something that can meaningfully be spoken of within a framework. Unicorns are not part of the framework of objects, and so they do not exist, but this is a trivial fact about the words we agree to use working within the framework. Finally, within the framework of propositions, 'electrons exist' is licensed by the rules of inference constitutive of that

framework (as long as the framework is supplemented with a framework that licenses talk of electrons).

Externally, the questions correspond to independent existence claims and the truth of sentences. But these just cannot be answered without the use of frameworks that govern the norms of proper use and misuse of a word – linguistic/conceptual frameworks are a precondition of meaningful utterances. So the question, "are unicorns *real*" or, more saliently, "do y/J particles *exist*?" when asked in an external sense is unanswerable, since the subjects of these sentences are literally meaningless independent of a linguistic framework. The only legitimate external questions are about whether it is *useful* or not to speak in accordance with a linguistic framework, and these can only be pragmatic: a choice of frameworks helps us adapt to and control the world, predict future experience, etc.

Returning to Fine, he doesn't explicitly endorse a Carnapian distinction between internal and external questions. But I do think there is a natural fit here, when NOA is suitably interpreted. Realism and anti-realism are best seen as attempts to answer the external questions of real existence and truth, or so I contend. I do not know that Fine would say, with Carnap, that the realist and anti-realist are literally asserting nonsense – perhaps we can understand what each is trying to do. Then again, a table thump has no semantic content – it is essentially performative: "Listen up! When I say electrons exist, I *really, really* mean it!" Either way, Fine does suggest that what each tries to add is in a

certain sense fruitless, and provides *nothing* over and above the existing scientific framework we find ourselves in that can help us make better sense of this framework. We should be happy to stick with "internal questions" of ordinary scientific practice; although the external questions might be meaningful (and this is still an open question), as departures from *his* minimal "core position," they *pragmatically add nothing* to the bare content of scientific theories. Thus a choice between them depends on one's antecedent emotional attachments – either pessimism about the pretensions of science, or optimism founded upon its remarkable explanatory and especially predictive power, instrumental value, etc.

I said above that offering an argument against the skeptic is not enough *without a further argument* to establish epistemological realism. But, if the reader will permit me to embark upon a brief aside, one current trend in philosophical responses to global skepticism (with which I sympathize) is to supply the missing argument. That is, there are *reasons* we can give to say why once skepticism has been discharged, we have no reason to doubt our ordinary knowledge of the world. But this is a *distinctly philosophical view*, since it is a reflective justification of our knowledge – as long as one says why the defeat of skepticism is enough to establish what is now an *explicitly held* epistemological realism. It is not the quietist despair that realism and anti-realism equally hopeless.

First, concerning global external world skepticism, some have developed "direct" theories of perception and have pointed out that the great mistake of British

empiricism is the Lockean "theory of ideas" that Hume adopted uncritically. This is the idea that the direct objects of perception are mental representations, or "ideas," but this seems to be committing a category error. Instead it seems that we should say that mental representations are the *vehicles through which* we form beliefs and representations about the state of the world – so it is wrong to say I perceive an *idea* of the laptop in front of me. Instead, what I *see* is a laptop. And I see this *by constructing a visual representation* of it – for how else could one see? This is not to say representations can't be mistaken: they can in all sorts of ways (perceptual illusions, low illumination, ingestion of psychedelic drugs, etc.). But in general, the fact that we have representations provides no reason to *doubt* that we interact with a real world. Mental representations are intermediaries between the world and ourselves, but they are the *means* through which we perceive things, not the objects of perception themselves.

Furthermore, the growing popularity of naturalistic epistemology points in takes the general form of Hume's solution, but without abstract philosophical worries about our second-order knowledge of the world.[5] According to these philosophers, some of whom identify as "neo-Mooreans," we can't help but start from within the starting point, and so it is just as legitimate in some sense to insist upon the knowledge we seem

---

[5] At the same time, there is, I think, a consensus that Hume's worries about personal identity and *especially* causation and induction in general remain unsolved, if not insoluble! There is, however, a vocal minority of opponents who think that we can recover the latter kind of knowledge by having a direct "impression" of force or necessary connection through *acting* on things in the world, such as the sensation of *being the necessary cause* of a ball's trajectory through space when we throw it. I won't comment on these arguments here.

to have, defeating skeptical hypotheses. Taking a different tack, Penelope Maddy has recently criticized philosophers such as Stroud who try to raise the spectre of Cartesian doubts (Maddy (2007)). She points out that the skeptic asks us to bracket *every single method we have* for evaluating a hypothesis, and then unreasonably demands us to show, once we are disarmed of all of our epistemological tools, to work back to what I have called the "starting point." Certainly, this has the air of triviality (perhaps even unfairness) about it – sure, I can't prove to you *without* using any of my methods of belief-formation that these methods are in general reliable, but so what? *What else should we expect?* The skeptic's threat, Maddy contends, is at least a *pragmatic tautology* – "if I can't perform *p*, then I can't perform *p*," where *p* is the task of justifying my epistemic methods *without* appealing to any of those methods. The skeptic can try to inflate this into a worry that we should take seriously, but like other tautologies, this fact is in a very real sense *devoid of non-logical content* and so shouldn't trouble us much.

The common thread that runs through naturalized epistemology is the idea that it was wrong in the first place to expect a transcendental method of fact-finding that can yield *certain* foundations for knowledge. Such foundations are of course unattainable, and the demand for them can seem silly once we are disabused of this idea. Instead, since we aren't logically omniscient beings, we *have* to "start in the middle" (from the "starting point") and build outwards from there, refining our methods as we go, making errors, but also learning from them as we go to arrive at *pragmatically certain* truths about

the world (e.g., that the Earth is flat, or maybe even that it is about 4.5 billion years old). The foray into naturalized epistemology is to mention, for the sake of comprehensiveness, that many philosophers do think that an adequate answer to, or dissolution of, the skeptic's arguments is enough to arrive at a second-order, reflective epistemological realism.

Applying this to anti-realism as a selectively skeptical hypothesis about our knowledge, this would be the position that defeating the argument that $\neg\mathbf{E}_{PP}$ and therefore (in the absence of good reasons) Epistemic Parity holds by default, now as a *philosophical*, *second-order* hypothesis. But it would be a mistake to leap from this to the thesis that we know everything that our best scientific theories claim, i.e., a very global version of $\mathbf{E}_{TK}$. For there may be good reasons to doubt, in specific cases, that a scientific theory tells us what the world is really like (again, consider the persistent problem of reconciling General Relativity and Quantum Mechanics, or the recent discovery of epigenetic factors in developmental biology, the details of which are still very much cutting-edge science and therefore should approach its conclusions with a degree of caution, with openness to the fact that we may revise our current opinions).

It would then be up to the realist to establish, on a case-by-case basis, that we *do* have scientific knowledge in certain cases. This would be a kind of second-order, *reflective* and *philosophical* analysis of what we know and not just ordinary scientific work – scientists do not, generally speaking, engage in second-order reflection on what we can

and cannot know on the basis of theory; this is the work of epistemology. But it is a call for epistemologists to work very close to the actual facts of scientific practice, and to articulate the underlying reasons why we should, say, accept the atomic theory of matter as about as certain as other forms of more mundane knowledge, but that we should only tentatively accept that dark matter/dark energy is a probable, but much less certain, hypothesis. In other words, this is a call for a rather extensive *localism* – attention to the concrete, contextual details of each specific case – in scientific realist arguments. There are deep, genuinely philosophical problems that emerge *within* scientific practice that mirror the very general, global arguments that philosophers of science have articulated – Lawrence Sklar (2000) has masterfully argued that this is the case in physics, showing how (say) reduction of a theory to its empirical content is called for in certain problems in Quantum Field theory, or that the notorious "hole argument" poses a *theory-specific* underdetermination problem. This preserves a unique role for the (epistemological) scientific realist without collapsing into an *anti-philosophical* quietism or non-realism (such as NOA) but one that is sensitive to context and averse to settling questions of realism in the abstract. Sklar has also forcefully and quite persuasively argued just this point – that this is a distinctly philosophical project, wed as it may be to the specific details of a theory. As a personal note, this is the kind of argument for realism I find most promising and intend to pursue in the future. But since I am not *arguing* for realism

here, I will restrict my comments to merely pointing out this possibility, as a signpost towards future inquiry.

Allow me to summarize. I have argued that the best sense we can make of the dispute between scientific realists and anti-realists is by analogy to the more general philosophical debate over skepticism and epistemological realism (the idea that we *do* have knowledge of an external world). On this view, scientific anti-realism is just *selective skepticism*: skeptical arguments that are specific to worries about science as a knowledge-seeking enterprise, leaving our mundane knowledge untouched. This requires demonstrating that there is *something different about science*, at least in local contexts, such that a Principle of Epistemic Parity fails to obtain, i.e., and argument that ¬**E**$_{PP}$. The anti-realist offers a positive reason to expect that theoretical knowledge can't be obtained, whereas "ordinary" knowledge can.

Non-realism is essentially an agnostic thesis: proponents do not deny **E**$_{PP}$, as this is constitutive of *anti-realism*, and might even affirm it, as Arthur Fine does.[6] But this does not mean that they go beyond this to think that we can *know* that we have some theoretical knowledge, as **E**$_{TK}$ would require. They adopt an anti-philosophical attitude towards such questions about whether we *really know* what we take ourselves to know,

[6] Just as scientific realism and anti-realism can be *local* theses (cf. below), such that one can take differing attitudes towards one or the other, non-realism as well. The best example of this is Turner (2007), who defends what he calls "The Natural *Historical* Attitude" – non-realism about assertions about the past, along with (it seems) scientific realism about the experimental sciences.

and so return to the epistemological starting point of uncritical, unreflective acceptance of knowledge claims without any philosophical justification of this idea.

Finally, scientific realism must be (cf. my further arguments below) the thesis that **E**$_{TK}$ is satisfied at least in some local contexts. It *is* consistent with local anti-realism as well: one might think that (to return to an example I have frequently used) physical theories are too transient to place our trust in, but that the central theoretical claims of chemistry and biology are even more secure than some cases of putative mundane knowledge, such as a detective's working hypothesis or very tentative conclusion that Smith rather than Jones is the arsonist. It is too weak to assert the mere possibility of knowledge without asserting the *actuality* of such knowledge, to deny or remain agnostic about **E**$_{TK}$. Such a position is equivalent to the claim that we are not in a position to know, at least currently, whether this we have any theoretical knowledge, even though nothing in principle prevents us from attaining this.

One way to see this clearly is to consider the main argument for realism, that *only* theoretical truth, or approximate truth (semantic realism) could explain the fantastic success of science. Although I find these arguments fundamentally problematic, one must still consider these arguments to examine *what it is scientific realists seem to be arguing for*. Belief in **E**$_{PP}$/ **A**$_{AK}$ combined with either ¬ **E**$_{TK}$ or mere agnosticism about **E**$_{TK}$ is obviously much weaker than what these arguments intend to establish. Supposing (as I think is reasonable) that the general "no miracles" argument for scientific realism that

has *absolutely dominated* the literature gives us a clue as to what scientific realists are *really out to establish*, this position is disqualified as a candidate for the label "scientific realism."

This is because these arguments generally conclude, "scientific theories are (approximately) true," on the basis that their success is otherwise inexplicable, which I will rephrase as "(some) scientific theories give us (approximate) knowledge of the world." Just saying "we *could*, in principle have some approximate knowledge of what the world is like" as an explanation of success – *even* in its sophisticated forms (e.g., separating out "idle" from "working posits" *per* Kitcher (2001), or the emphasis on novel prediction *per* Leplin (1997)) – is not enough to arrive at this conclusion. We need *actual knowledge* to explain what the scientific realist wants to explain. So I think it is fair to require that (epistemological) scientific realism be characterized as something much stronger, a positive acceptance of **E**$_{TK}$.

Here are the three options, expressed both globally and locally:

- **Scientific Anti-Realism:** For all/for certain scientific theories, an anti-realist believes ¬**E**$_{PP}$: It is not the case that all/certain theories *could* supply knowledge of what the theory literally says (e.g., think of constructive empiricism, social constructivism, or neo-instrumentalism).

- **Scientific Non-Realism:** For all/for certain scientific theories, a non-realist *does not believe* ¬**E**$_{PP}$, nor does she believe **E**$_{TK}$:[7] We should be agnostic as to whether all/certain theories yield knowledge (anti-realism is true) *and* as to whether all/certain theories do yield knowledge (scientific realism is true).

- **Scientific Realism**: For all/for certain scientific theories, a scientific realist believes **E**$_{TK}$ (and, *a fortiori*, **E**$_{PP}$): All/some theories can provide us with knowledge literally expressed by the theory (not merely of its "empirical content," or of its useful consequences), and all/some theories actually do provide us with this knowledge.

Stepping back from this argument and taking the broad view, although I find the analogy to skepticism to be helpful, and although I believe that it shows that we can capture everything we want about the scientific realist/anti-realist dialectic by

[7] One might be tempted to phrase this as ¬(¬**E**$_{PP}$ ∨ **E**$_{TK}$), and then apply De Morgan's law to this proposition to obtain (**E**$_{PP}$ ∧ ¬ **E**$_{TK}$). This doesn't seem right, though, for the non-realist is *agnostic* between the anti-realist's and the realist's view, but this conjunction asserts the outright *falsity* of both anti-realism and realism, that scientific knowledge is in principle possible, but in fact does not exist. This corresponds to the "anemic" position I surveyed above, which wasn't quite realist, but certainly "halfway there." The non-realist instead wants to say something like both the anti-realist and the realist are making a category error, or perhaps that neither position is knowable, and so on. So what went wrong? Well, **E**$_{PP}$ and **E**$_{TK}$ are *epistemic propositions*, and furthermore, the concept of negation, of falsity, is not the same thing as *non-belief*. We have wrongly equivocated between these two concepts. The appropriate logic to express the full content of this and the other sentences is a form of modal logic, epistemic logic. These are abbreviated as atomic propositions, even though they strictly speaking have a much more complex logical structure, and introducing the machinery necessary would be otiose and would overcomplicate the explanation. Fully written out in epistemic logic, an application of De Morgan's law *does not* assert that both scientific realism and anti-realism are downright false, but instead that the non-realist *does not believe* either.

characterizing the debate in epistemic terms, there remains a serious worry about considering this a "proof" of the priority of epistemology. Could we recast the debate in, say, metaphysical terms, or as a debate over how we ought to read the historical record? I don't think that is likely that we will find the task so easy. How could I prove such a negative claim, however? Furthermore, as I said, these considerations function primarily as appeals to intuition, rather strong intuitions, I believe, but this will only take us so far. It falls short of, say, the standards of deductive proof. I think there is, in fact, a much stronger case to be made that epistemology *has* to be the primary battleground over which the scientific realism debate is fought.

Recall that my overarching project is a search for a minimal "core" doctrine, acceptance of which is enough to be considered a scientific realist, what I labeled the "core realist commitment," or the **CRC**. It should be abundantly clear by now that I endorse the following claim: that the **CRC** = **E**$_{TK}$; that the Theoretical Knowledge Thesis is the right candidate for the core realist commitment. I view scientific anti-realism as a subspecies of philosophical skepticism, a negative epistemological stance, and scientific realism is the logical contrary of this claim, a positive *affirmation* of this knowledge. My arguments thus far are intended to prime this intuition, but I also believe an even stronger case is to be made when takes this bold epistemological claim *in isolation from as many other views as possible*, and sees just how far this assumption will take one without seeming to sacrifice anything that violates the central intuition underlying the "ultimate

argument" – that (some) current theories are not merely useful fictions or dispensable, but are ineliminable from our world-picture, and as such need to be *affirmed* (with qualifications where needed), without offering some special philosophical gloss that "explains away" its apparent real achievements at explaining the way the world is (to characterize things as neutrally as possible).

The much wider variety of positions that are consistent with the **CRC**/ **E**$_{TK}$ than we imagine there are, however, these aren't properly considered *versions of* scientific realism. That would lead us back into fragmentation and confusion about what scientific realism *is*. Rather, each of these possibilities should be considered optional *add-ons* to scientific realism that are *independently* motivated. That is why the Theoretical Knowledge Thesis is best seen as the ***core*** *commitment* of scientific realism: whatever else one believes must come from extraneous philosophical commitments, or *perhaps* (in special cases) from the content-specific details of the theory towards which we adopt a realist attitude (although we shouldn't expect these to hold in general). In a slogan: *Scientific realism, by itself, just is the bare thesis that we have at least some knowledge of the world that is disclosed* ***only*** *on the basis of a scientific theory, along whatever other absolutely minimal philosophical views this logically entails*

Some of the resulting [**E**$_{TK}$ $\wedge$ $\neg$**Φ**]'s may seem a little unorthodox and strange at first, but I argue that this is a "feature," not a "bug" (borrowing a turn of phrase from software programmers). Remember that the *very point* of finding the **CRC** is to *distill*

scientific realism down to its "essence" if you will. It should not be surprising that the result a very minimal, very ecumenical doctrine that is tolerant of a wide range of disagreement over unrelated philosophical views, held for other, independent reasons. And this broadened horizon of what is consistent with scientific realism is, intrinsically, a *good* thing. Much of what we take to be at odds with traditional ways of understanding scientific realism – the "prejudices that need to be dispelled" I mentioned above – are perfectly consistent with scientific realism, once we have an appreciation for a minimalist epistemological interpretation.

6. Excising Metaphysics from the CRC

First, let us consider just how far we can take the idea that we have robust claims to theoretical knowledge without requiring that is be knowledge of a conceptually independent domain of facts, as we discussed in the context of **M**CI. There I hinted at the prospect scientific realism is perfectly consistent with the possibility that *some* facts and phenomena are socially constructed. For example, this might be a reasonable attitude to take in some of the medical sciences. There are historical examples of diseases, which might be said to have been constructed from our beliefs and attitudes. Hysteria in women is one such example. On one reasonable way of looking at hysteria, the "disease" could be said to have real causal potency in the sense that women *labeled as* hysterical typically exhibited the symptoms associated with the disease, such as generalized anxiety, fainting, and so forth. Yet this is not because hysteria is a real disease of the mind (or body) in the way that other disease-processes such as cancer are "real." Hysteria itself, a mind-independent existing process, did not inflict these sorts of symptoms. Rather, according to this point of view, the categorization of women as hysterical subtly influenced the behavior of those so-categorized to conform to the stereotypes associated with the "disease." There is no reason to believe that we have eliminated socially constructed syndromes in the medical sciences (but of course this is not to say that we have not made progress in medicine).

Ian Hacking, in his (1999) study of social constructions has astutely described cases of socially constructed syndromes, like hysteria, as the product of our own behavior, as "looping effects." The collective decisions of a community to treat some category as describing a natural-kind or marking a real boundary in nature where no such natural-kind or boundary exists can paradoxically bring the category into being (in a socially constructed sense), such that items (or more typically, persons) that never before fit into the category, suddenly do. These are nothing less, and no mysterious than, other forms of self-fulfilling prophecy. Another salient example is the supposed link between race and intelligence. If an educator even subconsciously believes that, for instance, children of African-American descent are less apt at tasks like reading and arithmetic, this expectation may lead the educator to give such students less attention (since they are hopeless cases), or rate their work lower relative to children of other ethnicities. As a consequence, the stereotyped children devalue their own self-worth and academic talent, and therefore treat education as unimportant, and therefore score lower on standardized tests. The belief that these children are less intelligent could causally contribute to the eventuality that the children really are less intelligent.

Although there is clearly *causation* here of a pernicious sort between a community's collective belief and the behavior of some of that community's members, examples like these are not mere species of what I labeled "causal dependence" of either the original or sustaining sort. The difference between hysteria, or the looping effects of

race and intelligence, and things that causally depend upon our activities like ununoctium is that the very collective belief that hysteria is a real illness or that race makes a difference to intelligence is what causes these very things to exist. By way of contrast, ununoctium does not require that anyone have a concept of it in order to exist. It is metaphysically possible that a sample of it could be produced under very special conditions elsewhere in the universe, or as the by-product of an experiment by unwitting scientists who lack the very concept of a chemical element. Put simply, if we all stopped believing in hysteria (as the medical community eventually did), or abandoned the (conscious or unconscious) idea that the color of one's skin makes a difference to one's intelligence, these things would go away, would cease to have real effects. *This* is the essence of conceptual dependence: beliefs and attitudes that **X** is real *make* **X** real, and **X** wouldn't be real if no one had these beliefs and attitudes. Note that I am not claiming, and do not mean to imply, that socially constructed things are in any way *unreal* in the sense that they do not exist. Eating disorders such as anorexia nervosa are diseases of the mind and body that are arguably caused by cultural stereotypes of feminine beauty in Western culture. It is reportedly not found, or only very rarely found, in non-Western societies (Lee, 1995). Accordingly, some medical anthropologists have referred to it as a "culture-bound syndrome," an illness that is found only within a particular culture, but is unfamiliar to (and unfound in) outsiders. These syndromes present no objective bodily disorder in the sense of any tissue or organ damage, but they

do or can have a somatic component. No one would suggest that anorexia nervosa is unreal or that it does not have catastrophic effects on its victims. Anorexia has claimed lives, but unreal disorders cannot kill people. Yet, if this view of the disease is correct, it has its origin in purely social causes. Were we to have different standards of female beauty, the disease would (arguably) not exist. Again, we find a sort of conceptual dependence in this case. This example serves to show that we need to be careful to make a distinction between something's being unreal and it's being socially constructed. It is, I believe, reasonable to regard some disorders and syndromes like these as social constructions. However, these things that we construct are certainly real. It is just that they our counterfactually dependent on our having certain concepts; if we thought and behaved differently, these things would not be real.

Bringing the discussion back to scientific realism, the reality of some socially constructed categories introduces an interesting complication for the realism debate. What should the scientific realist attitudes be towards things like culture-bound syndromes and other socially constructed phenomena? In one sense, these things clearly exist. But in another sense, we are tempted to say that diseases like hysteria never existed in the first place. I think there is no contradiction here, appearances to the contrary. Hysteria is now defunct as a medical category; no certified medical practitioner who is in her right mind would use the category to describe and explain someone's behavior. As a society, we have “gotten over” belief in the reality of hysteria,

and this has robbed the category of its causal efficacy through the "looping effects" that once obtained. Thus, from our current perspective, hysteria never *really* existed, at least not in the sense of something that exists independently of what we think or say or do. The words "exists" and "real" are being used equivocally; hysteria exists as a conceptually dependent entity, but it does not exist as a mind-independent one. Now, insofar as scientific realism involves a search for things as they really are, "out there," we should be anti-realists about hysteria and other such things. But perhaps a more liberal scientific realism, one more closely allied with the human and social sciences, might be interested in what's real in this broader sense. In the interests of pluralism, see no good reason to foreclose this perspective on scientific realism, and a realism of this sort certainly has not been explored and articulated (as far as I know).

Localized versions of scientific realism lend themselves well to the kind of pluralism I am describing here. One may be an **M**CI-realist about quarks and lots of other unobservable and theoretical entities of physics in the robust mind independent sense of the term "realism." One could then be an **M**CI-anti-realist about hysteria, or other socially constructed syndromes in the very same sense. Consistently with this, one could also be a realist* about such things, where realism* about **X** is understood as the looser thesis that **X** exists, but only as the result of certain social processes. Nor need the looser sense of realism yield a bloated, Meinogian ontology according to which everything we have a word or concept for exists, if only in a socially constructed sense. There are

clearly many things we once thought to exist, but do not and have never existed even in this looser sense that allows for conceptual dependence. There is a very clear way in which something like hysteria existed because of the causal "looping effect" chains that affected people's behavior. Contrast this with something like phlogiston, a paradigmatic failed theoretical posit of early chemistry. There are strong constructivists who do want to claim that *everything* is socially constructed, including everything in the ontologies of both successful and rejected scientific theories. But this is an extreme position, and it requires a stronger burden of proof. The strong constructivist also owes us a plausible mechanism to explain how the social negotiations of scientists could make (formerly) non-existents real. We have an unproblematic explanatory schema for the types of social constructs I find unproblematic. In every example I've used, conscious or unconscious mental processes subtly influence behavior to make human beings conform to the stereotype associated with the socially constructed category. How do the decisions of the scientific community, and their social interactions, bring phlogiston into being (staying with our example)? In the unproblematic cases of social and medical sciences, people self-identify as belonging to a constructed category that never before had any members. *Human behavior*, specifically, conformity to socially prescribed norms and values, whether this behavior is conscious or unconscious, is the operative causal mechanisms for "making" social constructed things and persons. What is the analogous process for phlogiston? The explanation of how phlogiston had better be of a different kind. If we

try to shoe-horn phlogiston into the same explanatory schema for the unproblematic cases, the best I can do is to postulate that there exists some "stuff" out there that never was phlogiston before 1667 (when phlogiston theory was first articulated), but was responsive to scientist's theoretical beliefs and expectations, such that it *acted like* phlogiston for them, i.e., it took on the properties phlogiston was believed to have. Presumably, now that we have different theoretical beliefs and expectations, this "stuff" has shed its phlogiston-properties and taken on new ones. This is quite a curious metaphysics to adopt. A thing that conforms to our wishes and expectations would be a valuable thing indeed (scarcity of resources would be a thing of the past)! Now, I am well aware that this is not doing full justice to the strong constructivist, that their best arguments are quite more sophisticated. My point here is not to refute strong constructivism. It is to merely point out that there is an explanatory lacuna, and therefore a liberal sort of scientific realist who embraces the reality of conceptually dependent, socially constructed things and persons is not thereby committed to a much stronger view that everything is socially constructed. We need an independent argument for *that* claim.

Of course, matters aren't quite this simple. It cannot be social construction "all the way down." It cannot be that beliefs can fix the facts of the world, *not* via a perfectly intelligible mechanism of self-fulfilling prophecy and mild social construction that we see in the human sciences, but as if by magic – a kind of "Disney-World"-like

mechanism in which wishing somehow makes it so. Down that path leads the way to cognitive relativism and ultimately inescapable logical incoherence. When it comes to the objects and theories of the "hard" natural sciences, I am in full agreement with scientific realists of this particular stripe. But as I've tried to suggest, there's neglected interest in the social sciences concerning the "reality" or "existence" of various socially constructed phenomena which *are* real or could be said to exist in the sense that, while they are dependent upon, *created in virtue of*, what we do (say, think), still have real causal efficacy. To use a different example, gender differences between prototypical male and female behavior, especially behavior concerning role-preferences in domestic relationships, *may* be a social construct instead of genetically hard-wired. *If* gender differences are social constructs, they are nevertheless *real* and (potentially) have causal efficacy, through self-fulfilling prophecies, or Hacking's "looping effects." Parents who believe that female offspring are innately more likely to prefer passive roles in personal relationships, domestic life over a professional career, *et cetera*, may subtly influence their offspring's behavior and thus *bring about* the very differences that *wouldn't exist without* the sustaining/constitutive belief in the objective reality of these differences. A form of scientific realism which studies things that exist* in this sense (let us use the terms "**exist***" and "**real***" – adding on an asterisk – to denote things of this nature, which exist in this second, mind-*dependent* sense) has to this date been unarticulated, and is well deserving of attention. Our common-sense scientific realist should be

sensitive to this issue and therefore at least open to the possibility of such a "liberalized" scientific realism. She accepts **M**CI in its narrower sense for the natural sciences, but is open to (and perhaps excited about the possibilities for) rejecting **M**CI in certain cases within the social sciences, and allowing for the study of *real**, but still socially constructed, phenomena.

Moving on to **M**NK, at first glance it appears that it coheres quite well with a robust epistemology that stakes out some theoretical knowledge of the world. The reason for this is that the view that the world has a natural-kind structure (recall, this is just what **M**NK says) is a substantive metaphysical view about the way the world is. This is certainly true, and as I mentioned, Psillos doesn't need something quite this strong to defeat the anti-realist worry he advances the thesis against. However, I think that there is a more qualified sense in which **M**NK could be local. It could be that the world has more or less "pre-built" structure, or "joints" for science to carve out in different domains of nature, thus different scientific disciplines will differ in that some will have more natural-kind structure to discover, while others will need to rely more on conventional forms of classification. A different but equally relevant way in which natural-kind structure could come in degrees enters the picture when we consider levels of explanation or scientific description of the very same physical system. An example of two levels of description of the same phenomenon is the difference between microeconomics, which explains in terms of the decisions of individual people and firms

competing for limited resources, and macroeconomics, which explains at the level of the net effect of each of these individual decisions and uses indices that represent overall trends in the economy, such as unemployment and GDP. Let us take each of these cases one at a time.

In the case of totally different scientific disciplines, ones that study different types of phenomena, it just might be a contingent fact about nature that some objects fit more neatly than others into natural classification schemes. Spelled out using the analysis of natural kinds I offered above, this means that some object-types (like atoms of gold) or groups of object-types (all nomologically possible molecules) have an *essence*; a shared property or group of properties that serves a special explanatory and descriptive role. Essences explain why objects that share a particular essence share other more superficial properties, such as color, taste, ductility, conductivity, *et cetera*. They do this because there are laws of nature linking essential properties to the superficial ones. For instance, pure ethanol has a flash point (the lowest possible point at which it vaporizes to become a flammable gas) of 12.7ºC (incidentally, since this is significantly above the point at which water freezes, say, in cold winter months, this is why biofuels like ethanol must be mixed with more flammable substances, such as normal gasoline, in order to serve as a functional vehicular fuel). The flash point of 12.7ºC is a function of, or a nomological consequence of, the structure of ethanol molecules, as well as the binding energies of the covalent bonds between the constituent molecules. It takes a certain

amount of energy in the form of heat to break the existing covalent bonds in a molecule of ethanol so that the atoms recombine to form carbon-dioxide gas and water vapor. As ambient temperature is the a measure of the average mean kinetic energy of molecules in the surrounding environment (heat), this value for the flash point of ethanol is *entailed* by the relevant laws of nature and the structure of ethanol, given by its molecular formula C(2)H(5)OH. This structure is a plausible candidate for the essence-role because it nomologically necessitates many other such properties of the substance. Chemistry in particular lends itself well to natural-kind essentialism. In general, chemical structures of (pure samples of) substances nomologically determine many such superficial properties.

However, the hypothesis of natural kinds might not be as well suited to other scientific disciplines. In fact, chemistry seems to be a best-case scenario. It is much more controversial, for instance, whether there are natural kinds in biology. Specifically, within the philosophy of biology, it is contested whether biological species form natural kinds. The debate is complex and nuanced, and I cannot do it full justice with a brief summary. Suffice it to say that there is a range of positions to take, from realism about the species concept to species anti-realism. The debate is further complicated by the fact that scientists have articulated multiple ways to conceive of species. For example, are species best thought of as groups of interbreeding populations (Mayr 1942), or lineages of individuals with a common ancestor (Hennig 1966). It is commonly agreed that there is no *one* species concept, and that the multiple concepts cross-cut in the sense that an

organism can be classified as belonging to one species according to one concept, and to another on the basis of another concept. Some authors take this pluralism to be at least consistent with a mild realism about species (for instance, Dupre 1993), while others take it to clearly imply anti-realism (Ereshefsky 1998). Suffice it to say that the fact that one can "carve up" organisms into different categories seems to imply that species have no shared essence in the same way chemical substances do. According to a kind of epistemological realist pluralism, one can individuate species in different ways by identifying a shared essence that will vary with the operative species concept. Again, for example, viewed as interbreeding populations, a species' shared essence is the fact that any two opposite-sex fertile members of the species can reproduce fertile offspring, whereas viewed phylogenetically, species' have a "historical" essence consisting of the fact that all members of the species are descended from a single common individual. One reason for embracing anti-realism follows from embracing a strong form of **M**NK, according to which there is a *unique* natural kind structure of the world, a One True Story. If we assume this, the fact that there's no one common individuating feature that sorts organisms into one, and only one, taxon, a feature which explains all the superficial commonalities of these organisms, anti-realism follows. I do not wish to enter a judgment here over whether species mark mind-independent, real differences between biological types, or whether the kinds we sort organisms into are purely conventional –

for a version of conventionalism that is at the same time, in my view, mildly realist, see La Porte's excellent book (2004).

The central point I would like to make is that there is certainly less structure "pre-built" into the world when we compare that part of it biology studies to that of chemistry. If a mild pluralistic realism about species is defensible, if (as I believe) it is reasonable to think that there are *objectively* better and worse ways to group organisms, but *also* that there are multiple ways to do so, and perhaps no one way is better than any other, *then* it is *both* true that natural-kind structure guides our formation and use of species-concepts, *but* that the objectively existing structure of nature provides less "signposts," "joints," or *guides* to organizing species than it does chemical elements. If what I've termed "mild pluralistic species realism" is tenable, one consequence is that the ways to organize (some of) the things that biology studies (species) are *less* fixed by *the way the world is* than are chemical elements; that there are more ways to "carve up" species than there are chemical substances. So there can be variation in the degree to which the world is structured into natural kinds *between* different disciplines.

Turning to our second example, consider sciences that study the same phenomenon at different "levels" of explanation. Examples are not limited to economics. For instance, one often hears physicists speak of the "classical domain," where Newton's laws are operative (or at least empirically adequate). The classical domain is contrasted with the quantum world, which applies to micro-phenomena on small space-time and

mass-energy scales, and General Relativity, which studies phenomena on precisely the opposite extreme scales. Bohr's famous "correspondence rule" requires that theories agree with classical predictions in the limit where these (small or large) scales approach those with which with we are more familiar. Let's take the simple example of mass as it appears in Newtonian mechanics and Special Relativity. Newton identified mass as that feature of material bodies, which explains its acceleration in response to the net force on the body, as well as the property whose magnitude explains the gravitational force exerted on distant bodies. With the advent of Special Relativity, Einstein introduced the famous distinction between rest mass and relativistic mass. The latter is the mass of a body in inertial motion with respect to one's frame of reference. The former is the mass of the body as measured in a frame of reference where the body is at rest. This example raises some interesting questions: supposing that Einstein's distinction between rest and relativistic mass tracks a real kind-distinction, what should we say about mass as defined in Newton's mechanics? Note that here we are dealing with a property-kind (it might be even more precise to say that mass is a kind of physical magnitude) rather than object-kinds, but this is not (in my view) a relevant difference. It seems there are two views to take, parallel with the biological example discussed above, one anti-realist and one realist. We can either say that Newtonian mass missed out on a real distinction or boundary between natural categories, or we can opt for the realist line that Newtonian mass marks a real property that exists and has real causal efficacy on a coarser-grained

analysis of nature, in the "classical domain." For the anti-realist, there is not much more to say. Newton thought he found a natural-kind, but he was wrong. On the realist, "emergentist" type of view, we have to admit that Nature is structured into a hierarchy of "levels" in order to account for how the mass-concepts in both theories could track the real structure of the world. On the scale of higher relative velocities, there are two operative mass-concepts that track the real structure of the universe, but on a lower scale, one mass-concept serves to track some coarser-grained, but still "equally real," property. Levels of nature can be organized by differences we see in types of phenomena that "emerge" as we move scales of different physical magnitudes (thus I classify this view as a kind of emergentism).

Emergence of new phenomena, and for our present purposes, the emergence of new natural kinds, has both epistemological and metaphysical forms. Epistemological emergentism is the view that higher-level phenomena are not predictable or explainable from a lower level due to the computational intractability of making such predictions or explanations using only the resources of the lower-level theory. This intractability might further just result from the fact that we are just not smart enough and have not developed sufficient technology to perform the requisite calculations. The more interesting possibility is that the required calculations cannot be performed *in principle*, that even if we were armed with all the occurrent facts at the lower level, and given enough time and computational resources, we would not be able to predict or explain

what happens at the higher level. My example of mass is clearly not a good example of *strong emergence* in either an epistemological or a metaphysical sense, since there are "smooth reductions" linking the classical domain to both quantum and relativistic domains. (But even though we know how classical mass relates to relativistic and rest mass, the question remains of whether it is a real property; emergence comes into play when we think of nature as being structured into levels of different types of things, and I say classical mass is only weakly emergent because we have a grasp on how it does in fact "emerge" from the relativistic limit). A better example of strong epistemological emergence (strong in the sense that we can't individuate or identify higher-level kinds, or predict and explain their behavior, by confining ourselves to the resources of lower-levels) comes from Fodor's classic (1974). Higher-level kinds, such as currency, can have an infinite number of microphysical realizers. One token of the kind currency are pieces of paper in my wallet and various round pieces of metal in my pocket, all with the right causal history of being minted at one of the Federal Reserve banks. Another token might be information stored in my bank's computer chips, and there is no reason in principle why we cannot multiply examples, why we could not invent other forms of currency. To reduce the higher-level kinds to lower-level ones, we would have to embrace bridge-laws linking a higher-level type to an infinite disjunction of lower-level physical states. As Fodor famously pointed out, this is a strong reason for preferring anti-reductionism and regarding the special sciences as being "autonomous."

No strong epistemological emergentist denies that higher-level natural kinds supervene on microphysics. It is just that they cannot be identified, nor used to predict and explain, at this level. The metaphysical emergentist embraces a "spookier" view. Nature is indeed structured into layers of reality, and kinds of things emerge at higher levels that indeed cannot be predicted or explained from below. But this is not *merely* owing to any in-principle epistemological or computational limit. It is just a brute fact that there are levels of reality, and different types of things found at each level. On this view, emergent phenomena are something more than assemblies of micro-particles. It is very hard to express this view without using platitudes, such as to say that a higher-level kind like currency, or a cold-front is a "whole greater than the sum of its parts," something above and over the mereological sum of tiny things that are (what we wrongly take to be) fundamental.

The best justice I can do to this view, metaphysical emergentism, is to borrow a wonderfully apt conceptual metaphor from Jonathan Schaffer (2007). Ask yourself: what would it take for God to create the actual universe and everything contained within it? Both the reductionist and the strong epistemological emergentist agree at least the same answer, which is that God would have to do nothing more than to create all the fundamental particles in the right locations with the right various properties, i.e., to create the bottom-level. All the higher-levels come as an "ontological free lunch." The reductionist says we can make type-type theoretical identifications between lower- and

higher- level kinds, the strong epistemological emergentist thinks we can't do this, not even in principle, and since the best we can get are an open-ended disjunction of type-token identities, we should treat the higher-level phenomena as genuinely (strongly epistemologically) emergent. But on what it takes to make a universe, the two parties are in perfect agreement. Someone who prefers metaphysical emergentism thinks that God's work isn't done when He says "let there be all the fundamental constituents of reality." A metaphysical emergentist about cold fronts, someone who thinks that cold fronts are a *metaphysically emergent* feature of reality, thinks that in addition to putting a quark here, an electron there, and so forth, God has to also make the cold front (and only then can He take a day of rest!). I hope that this metaphor does justice to the position. After all, the metaphysical emergentist thinks that even the comparatively weak notion of supervenience that ties emergent phenomena to lower levels of reality makes the phenomena *too* dependent upon the subvenient bases which are their physical realizers. As a logical consequence, fixing the basement-level facts isn't enough to get metaphysical emergence.

It is worth connecting the two different notions of how phenomena could be emergent to **M**ᴄɪ, the thesis that the objects studied by science are conceptually independent of human activities. Metaphysical emergentism clearly secures a robust form of mind-independence, for emergent phenomena are *ontologically independent* of a subvenient base of fundamental physical facts. And I believe strong ontological

independence of this kind logically implies mind-independence. My intuition here is that if something like markets or cold fronts are not supervenient on lower levels of reality, then certainly they must be independent of our concepts, language, and practices also. Anything robust enough to enjoy this sort of independence must not also depend on there being human beings to conceptualize such things. If God couldn't create a cold front by creating a fundamental level, how could humans perform such a creative act?

Things are more complicated with the logically weaker belief in epistemologically emergent phenomena. The qualifier "epistemological" seems to suggest that without human beings creating and possessing concepts of things that are said to be emergent, such things would not exist. I think this inference is wrong-headed, and here's why: *ontological independence* from a subvenient base (the sort of independence metaphysically emergent phenomena enjoy, if there are any) is one thing, and *conceptual independence* from facts about human beings is another. We should not confuse the two. An anti-reductionist may firmly believe that there are strong epistemologically emergent phenomena, say, cold fronts. Our anti-reductionist will claim that even armed with all the facts about the fundamental level and logical omniscience, one still couldn't predict that cold fronts exist and how they behave. This belief is consistent with the fact that cold fronts wouldn't exist without there being fundamental facts for them to supervene on. Now, the existence and behavior of cold fronts may not be ontologically independent of some further set of facts, but we need an

additional premise to get from ontological dependence to conceptual dependence. (Note that against metaphysical emergentism I argued from ontological independence to conceptual independence. The argument forms are different). It seems obvious to me, in fact, that high-level phenomena like cold fronts and ecosystems exist independently of human concepts, language, and practices. I have to again consult the reader's intuitions, but isn't it the case that weather patterns and interconnected relationships between biological organisms and the environment would still exist were it the case if all human beings vanished from the face of the earth, one second from the present moment? Facts like these have a "Moorean" certainty in the sense that no person of common sense could deny that they are true unless they are under the sway of some philosophical theory that would entail such extreme skepticism. I defined "conceptual independence" in exactly this way, that the existence of human beings and their activities are what sustain the thing in question, not in a causal sense of producing something that *could* exist in a possible world where there are no humans, but in the sense of being *constituted by*, or *only possibly existing* in virtue of the existence of such human practices. These things (cold fronts, ecosystems) exist in a robust, "scientifically operational" sense that they license predictions and counterfactual claims, which themselves are perhaps instances of (*ceteris paribus*) laws of nature.

There might be other strongly epistemological emergent phenomena for which we need to offer a more careful analysis, such as markets and political trends. These

clearly would still exist, were humans to vanish in just the way described above. Does that make them conceptually dependent upon human behavior? Again, I would answer in the negative. Markets, for instance, depend entirely on the social behavior of human collectivities, but this dependence is not conceptual. Markets for grain (for instance) in human society presumably existed before the *concept of a market ever existed* or was articulated. The dependence here is *causal dependence* of both the *original* and *sustaining* forms. Our activities cause markets and political trends to exist, but such things *could exist* without someone having a concept of the things.

In conclusion, we have seen many senses in which we can localize metaphysical dependence on human behavior of both the conceptual and causal kind. There is a stricter sense of scientific realism in which socially constructed things like culture-bound syndromes just don't really exist. This view has some genuine plausibility to recommend it. These things aren't things that really are things *out there*, part of the pre-existing structure of the universe, the way Nature is organized (note: *not* the way *we* organize Nature. The two diverge). There is, however, a more liberal form of scientific realism that treats socially constructed phenomena as genuinely worth study. After all, they do have properties and characteristic behavior (hysteria has a stereotype of female behavior associated with it). So far as I know, scientific realists have restricted themselves to the stricter form. But realism about social constructions is an available position (at least, I believe this is so), and it is a position that has not attracted much attention, if any.

To get more specific, one robust way in which the mind-independent world might exist would be if it were organized into natural kinds of objects, with (to borrow Locke's terminology) real instead of nominal essences. As I have tried to stress, natural kind structure can come in degrees. I see absolutely no reason why scientific posits (whether unobservable, theoretical, or both) must *either* form a real category, a natural kind, *or* be a mere conventional grouping. There are better and less ways to organize nature, from a system based on a single set of properties acting as a robust "real essence," from systems based on clusters of shared important properties (such as Boyd's (1999) "homeostatic property clusters," kinds with real definitions held together by one or more mechanisms that explain kind-membership and surface similarities, but which may have fuzzy boundaries and bear only "family resemblances" to each other), all the way to completely conventional groupings based completely on nothing else than pragmatic interests. As I see it, the natural-kind structure of the world is a more-or-less affair. As consequence of this, I think it open to empirical investigation to find out just how much of the world, and to what extent, it is structured into natural kinds. This is a task for science itself, not armchair philosophy (this is a doctrine I will revisit later and further defend).

These differences exist across disciplinary boundaries, both between sciences that study completely different things, and between sciences that take two different perspectives on the same domain, from different "levels" of explanation and prediction.

I leave it as an open, empirical question just how much kind-structure is to be found in each type of scientific perspective on the world. And this attitude I'm expressing is a completely local approach to **M**$_{NK}$. And as I've argued, the kinds found at (or emerge from) different "levels" of reality, if there are any, are *not* conceptually dependent on human activity, at least not the great majority of them. There may be causal dependence, but this is a different matter entirely. This is an interesting way in which **M**$_{CI}$ links up with **M**$_{NK}$, notice in a local way. In sociology and perhaps psychology, we may expect to find a fair amount of socially constructed objects, things not existing independently of our concepts, which could not therefore be shoehorned into a "real" classification scheme, with no guiding "dictates of nature." In other, "hard," natural sciences like physics and chemistry, we might expect to find a one unique way to categorize nature that gets the real distinctions between objects correct. And, to reiterate for the final time, we might legitimately expect to find a continuum of natural-kind structure, where the world is completely "objectually inarticulate"[1] in some domains (pure conventional kinds), totally "objectually articulate" in others (objective natural-kinds with unique real definitions), or "objectually articulate" in different ways when viewed from different scientific perspectives (mild pluralistic realism, with potentially cross-cutting but still *natural* kind-structure).

[1] I borrow this wonderful turn of phrase from Iris Einheuser (2011).

Lastly in metaphysics, concerning **M**FT, the Factuality of Scientific Theories Thesis, the Theoretical Knowledge Thesis aligns with it in a rather straightforward and trivial way, such that there is no question here that epistemology will involve some *minimal* metaphysical commitments. Per the strong epistemological thesis, **E**TK, science informs us as to *what exactly* we can be confident about regarding our scientific image of the world, about what features of this world-picture are accurate. Up till now, this point has been made on the plane of epistemology. But it carries along with it some metaphysical, ontological commitments. *If* one trusts the scientific claims one puts their confidence in, if one in fact does *know* theoretical facts, then of course these unobservable things have to exist (and be very much like the way science describes them), and these theoretical explanations of why things happen have to be correct. I am not here making the mistake of drawing a metaphysical conclusion from an epistemic premise, not in any substantive sense anyway, but merely pointing out the metaphysical implications of what the world is like, *if* scientific realism is construed as the thesis that we have genuine theoretical knowledge.

7. Excising Semantics from the CRC

Simplistic or general statements of scientific realism generally seem to offer the platitude "scientific theories are true, or at least approximately true." Now, my proposal, that we try insofar as is possible to consider *knowledge* as something that is completely separable from *truth* must, *prima facie*, seem puzzling, or even incoherent. It is so often taken for granted that the mental representations which are the bearers of knowledge could only be *beliefs*, and those beliefs represent the world by being true or false of it in roughly the same way a sentence is true or false (i.e., sentences are like little sentences "in the head"), although there is absolutely *no* consensus on what turns a true belief into knowledge *besides* justification (I understand justification to consist of experiences, memory, reliable testimony, and other beliefs that are corroborating reasons). It is also overwhelmingly the default view that scientific knowledge of the world is just like more ordinary factual claims *qua* its mode of representing the world (as long as we are talking about knowledge and not *justification*, this is completely consistent with Epistemic Parity) I don't *disagree* with this picture of knowledge *per se*; that it is held in such high esteem indicates just how much philosophical work it can do. However, my overarching project is to restore epistemology to pride of place in debates over scientific realism, which requires excising as much as I can from the core Theoretical Knowledge Thesis. I don't think one is *forced* to accept such a view of knowledge, and the reason I do not simply take this picture for granted is that there *are* other ways to "be a scientific realist"

that reject the idea of truth as the right mind-world representational relation. This is not some highly theoretical, abstract, hypothetical view, but one that some philosophers actually accept![1] The proper place to actually argue for this will be in my discussion of the logical independence of epistemology from correspondence (and other) theories of truth.

My overarching concern is to shift the question of scientific realism away from philosophical territory where the debate has been (and is currently being) conducted, over the wrong issues. Crucially, I believe scientific realists are far too focused on questions in the philosophy of language and that their efforts would be better spent elsewhere, within epistemology and metaphysics. As I claimed at the very beginning, these are the questions we should be asking: Can we trust what science tells us about what the world is really like? Why? What current scientific theories should we take to be reliable guides to the way the world is, and which need more supporting evidence before we can place such trust in them?

At this point, a very obvious objection to make is that metaphysics and semantics are inextricably linked, for when it is the case that **P**, then any sentence expressing the content that **P** obtains is true. Now, perhaps it is a truism that the following statement: "sentence (or theory) **P** accurately describes what the world is really like" is synonymous with the statement "sentence (or theory) **P** is *true*." Perhaps this truism

---

[1] This is grist for my mill, that the epistemological version of realism is the most tolerant, minimal core, which is robust enough to vindicate an optimistic attitude towards at least some of our scientific theories.

even has the status of an analytic truth, if there are such things. I imagine that our common-sense scientific realist is bound to be confused at this point: isn't this just what we mean when we call a theory or an existence claim "true"? In an everyday, non-philosophical sense, yes, I am inclined to agree. "Representational accuracy," "an account of what things really are like" and "*truth*" are all probably equivalent and equally good ways of saying the same thing. I demur from using the language of truth for two reasons:

The very concept of truth is the focus of sustained attention, and is seen to be in need of an *analysis* or *theory* of what it is we do when we label a sentence or even, maybe, a thought (*qua* mental event), *true*. The correspondence analysis of the truth predicate that scientific realists are often so fond of attributing to theories is certainly not the only theory on the market. I have already mentioned the deflationary or minimalist theory of truth alone as one serious contender, and it would seem that one could get along just fine as a scientific realist without making robust commitments to abstract entities like propositions and word-world correspondence (i.e., the causal theory of reference). I do not believe, and am not here claiming, that the questions and problems that theorists of truth aim to address are in any way uninteresting or unimportant. But I do think that they have little relevance for the scientific realism debate, in the sense that the scientific realist has a very broad spectrum of options to choose from among these various competing theories of truth. There *are* theories of truth that are inconsistent with

scientific realism, i.e., Putnam's epistemic theory of truth and his attendant view of "internal realism," but this is due more to their metaphysical underpinnings or motivations which would conflict with a conception of ourselves as having objective knowledge of facts more than anything else that disqualifies them. So once again, I contend that it is epistemology, and not semantics, that is primary in this debate.

In the spirit of casting as wide a net as possible, I aim to make room for those philosophers who, while considering themselves scientific realists (perhaps of a peculiar or off-beat sort), are doubtful of the usefulness or even the legitimacy of the concept of truth itself, for whatever reason, *or*, those who are suspicious that theories represent the world by some means *other than by being true*, that the *essentially linguistic* model of a sentence's being true or false is the *wrong* model for how scientific theories represent the world. I have Paul Churchland in mind (and other self-identifying eliminative materialists) as a representative of the former type of scientific realist, and Ronald Giere as a representative of the latter sort. I will visit their views momentarily, but first I want to analyze truth-minimalism or deflationism. In the course of doing so, we will see how the perceived needs for a robust theory of truth like the correspondence theory, reference relations grounded in causal linkages stretching back to *natural kinds*, and an account of "approximate truth" to finesse the puzzle any semantic theory will face when it runs up against the fact that every scientific theory is necessarily incomplete, are

optional. They may be good solutions to the puzzles that *scientific representation* forces upon us (speaking personally, I incline towards these traditional views), but they aren't the *only ones* "on the table," and therefore they are not properly speaking required by any version of scientific realism. As before, this opens up room for non-standard versions of realism but which still capture the core commitment.

Why not endorse a correspondence theory of truth? For one, there are other theories of truth on offer in the marketplace of ideas that seem to have no negative consequences for scientific realism; they emphatically *do not* restrict or water-down, in any way, the (robust) knowledge-claims and ontological commitments that our common-sense scientific realist embraces. The correspondence theory, recall, has it that what makes a proposition *true* is a kind of "match" between word-and-world. The "direction of fit," the explanatory order, is world-to-word. The relation connecting terms and predicates (words, specifically, common nouns) to things and properties is that of reference, which, generally speaking, is standardly assumed to be the causal theory, often one that *requires* real differences between kinds to fix the extensions of terms. I have already explained these theories at length; so now let me discuss why I think they are extraneous.

The correspondence theory is a (metaphysically) robust and very detailed account of how language and perhaps thoughts are true. But is this account needed? What relevance does this have for the realism debate? The primary motivation, as

evinced in Psillos (1999, xx-xxi), seems to be that the correspondence theory is thought to be necessary to "warn off" a particular understanding of truth that is definitely inconsistent with scientific realism, epistemic theories of truth. We'll revisit these in just a second. But to show the irrelevance of the correspondence theory to the scientific realism debate, let's consider a third alternative: deflationary theories. In recent history, deflationary theories of truth have come to rival correspondence theories in popularity, and while they are not yet (as I understand the present situation) the dominant view, they seem to have displaced coherence theories of truth as the leading minority philosophical theory of truth.

Like "the" correspondence theory, there are many subtle details and differing versions of deflationary theories of truth, so my presentation of it is going to be sketchy and minimalist. The basic idea, however, is that truth is *not* a substantive property of sentences. The simplest way to express this thought is by way of the *redundancy theory of truth.* The redundancy theory has it that to assert the sentence "**P** is true" is simply to register one's assent to **P**, to say nothing over and above "**P**." Let's take a moment and consider what implications this would have for the scientific realism debate, if any. Let's compare the sentence "'electrons possess a spin of ½ is true," understood on *both* a correspondence *and* a deflationary account.

The correspondence theory of course involves a (comparatively) long and complicated story, involving facts about how terms like "electron" and "spin" came to

refer to (putatively) existing unobservable, theoretical natural kinds of things and properties, and that the sentence "electrons possess a spin of ½" is *made true* by the structured fact that all objects belonging to the natural-kind named by the common noun "electron" possess or exemplify the natural-kind *determinable* property named by the theoretical predicate "spin," and that the determinate value of this determinable property is ½ for all things belonging to the natural-kind named by the common noun "electron" (if one prefers to add, per the Natural Kind Hypothesis, this is a *kind-constitutive property* of all electrons. Phew).

Now, what does the redundancy theory have to say about the truth of the sentence in question? On that account, when I say "electrons possess a spin of ½ is *true*," what I am really saying is, "electrons possess a spin of ½." Now, if I am being *sincere* when I assert this sentence, if I am not lying to you, I *really do* have to believe that there *really are* such things as electrons, and that there is *really* such a thing as particle spin, and that the value of spin for all electrons is ½. To be sure, the correspondence theory answers a lot of facts that the redundancy theory cannot, or does not, (such as how words like "electron" and "spin" get to name the things that they do in fact name. But there is a certain simplicity and comparatively sparse ontology to the redundancy theory, towards which philosophers inclined towards such simplicity might incline. But here I aim not to compare the two theories as they stand on their own, to determine which theory we should prefer. My goal is much more modest. I ask: does our choice

between either account *really* make a difference to the scientific realism debate, to the idea that we are to take seriously at least some of the claims of science at face value? Is a correspondence theory over a deflationary one as an important issue that one needs to decide upon before formulating a coherent, substantive version of scientific realism? Yet again I consult the reader's intuitions. My personal belief is that she will view this as a debate that is internal to the philosophy of language, and one that has no bearing at all on the fundamental questions I raised above. Both the deflationist and the correspondence theorists (who are also scientific realists) agree that *there are* electrons, that *there is* such a thing as spin, and that electrons *really do* have spin ½. And they agree on virtually every other such scientific fact that we should regard as a safe bet. Their differences lie over what it is we mean when we attribute "truth" to a theory (conceived of as a set of sentences – but see my comments on Giere below). As my avoidance of the language of truth has (I hope) modeled, we can get by without even formulating scientific realism as a debate over whether scientific theories are *true*.

As I have stressed over and over again, what is *primary* in the debate over scientific realism is the epistemology of scientific claims, with the attendant metaphysical commitments it carries (but no "extra" metaphysics). I can best describe this by posing the question: "does science tell us what the world is really like?" My readers may find it absolutely shocking that I think we can get by without using the semantic notion of truth to characterize scientific realism, preferring instead to get to

(what I regard as) the heart of the matter; metaphysics and epistemology. Fundamentally, we scientific realists want to know if world really is the way science describes it as being, if it really is populated with the unobservable causal agents it posits, and if it really is governed by such and such causal laws, and *why* we can trust these claims, and regard them as secure instances of *knowledge*.

This all being said, I have to issue a strong cautionary note to the reader: not all theories of truth are equally welcome within the **CRC**. I am most concerned to exclude an epistemic theory of truth, which motivated certain realists' insistence on the correspondence theory in the first place. The problem with epistemic theories of truth is *their metaphysical underpinnings and presuppositions*. Underlying epistemic theories of truth is the thought that there cannot be, even in principle, a gap between what we could possibly know in an ideal completed science and what the world is really like. This view is a species of neo-Kantianism, and it has it that our concepts and perhaps our language structure the world into categories, rather than the reverse direction (the direction of fit is mind-to-world, rather than world-to-mind). This was to solve certain puzzles Hilary Putnam and those who followed him (the internal realists) found to be deeply troubling, such as how our concepts and language appear to succeed at dividing objects up into natural (but not objectively pre-existing) categories. His solution, recall, was that our language and concepts *create* these categories, rather than *tracking* the natural boundaries between kinds of objects (at least, this is how I understand internal realism).

(The layer-cake metaphor may be a helpful reminder to the reader, that the slice goes from top-to-bottom, where the top layer represents our conceptual scheme, and the bottom layer represents the natural-kind structure of the world. The two align only because the former determines the latter). The common-sense scientific realist by and large rejects this metaphysical scheme, although, as mentioned, she is amenable to constructivist hypotheses in certain social sciences, at least where such hypotheses pull their weight and do real explanatory work.

Notice again, however, that this debate over internal realism is occurring at the level of metaphysics and epistemology, concerning how our conceptual scheme and the structure of the world intersect. Putnam readily admits that he accepts an epistemic theory of truth, according to which a sentence is true if and only if it is warrantedly assertible from the perspective of an ideal, completed total science, *because* of his concerns (which I do not share) about how our minds and our language link up with the world. Yet again, this is more evidence that semantics is of secondary importance in the realism debate. Epistemic theories of truth are objectionable not so much on their own terms, because they get the definition of truth wrong, but because of the metaphysical scheme they presuppose, one which precludes the possibility of a meaningful scientific realism.

Finally, there is the second issue I raised above, the possibility that one could be a heterodox scientific realist, rejecting the realist platitude "theories are true," by taking

one of two routes: (1) One could reject *truth itself* as serving any useful or important explanatory role within the philosophy of language or elsewhere, or (2) One could believe that the linguistic notion of truth is a poor model for how representation occurs in science, that it is wrongheaded to think that theories can be called "true" in the same way that an ordinary sentence like "grass is green" is true.

The first view owes to Paul Churchland (1981), who, for quite independent philosophical reasons, rejects the dominant "research programme" within the philosophy of mind, which I will term "folk-psychological realism." According to this view, ordinary language propositional attitude ascriptions such as "Nathan *believes that* grass is green," "Mary *hopes that* ice cream is served for dessert," do *not* individuate and pick out event-types, specifically, natural-kinds of mental processes, or objectively existing brain-states, i.e., instances of, or particular acts and episodes of, believing or desiring. Churchland has a battery of arguments for thinking that our intuitive folk-psychological notions like belief and desire won't smoothly reduce to neurological states of the brain, and will instead be "eliminated." Let me clarify what is meant by "elimination" of theoretical posits: in the history of science, we find that occurrences of theory change fall along a continuum from smooth reduction to radical, disruptive change. One famous example involves the reduction of the phenomenological laws of classical thermodynamics of the mid-19th Century (as articulated by thinkers such as Sadi Carnot and Rudolf Clausius, the discoverer of entropy), to Statistical Mechanics. In

this case the former Laws, which govern the macroscopic behavior of thermodynamical systems, have clear microscopic "underpinnings." The classical laws are mathematically deducible from the statistical properties of large arrangements of tiny particles moving about and colliding with each other (later confirmed to be atoms and molecules), each of which has a variable amount of kinetic energy. Energy in the form of heat, earlier thought to be made up of an invisible fluid ("caloric"), was rather shown to be nothing over and above the kinetic energy of the constituent particles of a given physical system. Therefore temperature, previously believed to measure the "amount of caloric," was found instead to measure the mean kinetic energy of the system's constituent particles (where particle motions are relative to the center of mass of the system).

Actually, this is a nice example, for it illustrates smooth reduction in one sense. The kinetic theory of heat, which had been the minority view in the earlier half of the 19th Century, described heat as the motion of microscopic things too small to see. This insight was vindicated, for it turned out that the phenomenological laws of thermodynamics smoothly reduced to the "true" laws of thermodynamics, which, again, involve the statistical properties of large systems of interacting particles. The phenomenological laws, again, follow in a rather elegant fashion from what we can mathematically demonstrate about how such systems are overwhelmingly likely to behave, in a *statistical* sense (e.g., decreases in the entropy of a closed system have a small, non-zero chance probability of occurring, but it is a near-certainty that this will

almost never happen). These theorist's intuitions were vindicated: heat really did turn out to be the aggregate kinetic energy of the tiny bits of matter that make up a system.

Now, for the caloric theorists, this was a stunning defeat; for they conceived of heat as a substance, and invisible fluid, and that the phenomenon of heat transfer from hot to cold bodies was due to the fact that particles of caloric *repelled* each other, thus causing the substance to have an innate tendency to "spread out" in space. (Incidentally, this was taken to explain much else besides, such as why heated material bodies expand: as caloric particles enter the body, their self-repellant properties cause the body to expand from the inside-out). After the vindication of the kinetic theory of heat with Maxwell's and Boltzmann's work on statistical mechanics, "caloric" as a theoretical term was dropped and simply vanished from scientific language. *Here* we have a case of *radical theory change* with an attendant *elimination of theoretical terms* (i.e., "caloric"). Scientists found that there *just was* nothing even remotely close to caloric, as it was described by its proponents. Subsequently, we concluded that there is no such thing, that caloric does not exist. This is one example of *eliminative* or radical theory change in the philosophy of science, and we can list many other examples. Among Churchland's favorites are the posits of crystalline spheres perceived necessary to carry the planets about on their orbits, the invocation of demonic possession to explain schizotypal behavior and other forms of mental illness, and brute folk superstitions; beliefs in things

like witches and the evil eye. All of these things, by our current lights, just don't exist. They were *eliminated*.

Churchland emphasizes examples of these last types for the very reason that they were never part of any serious scientific theoretical framework, but instead the posits of *folk theories*; bodies of largely implicit, taken-for-granted, explanations of how and why things happen which fail to meet the standards that we typically hold modern scientific theories to, such as: an emphasis experimental confirmation, mathematical rigor, testability and falsifiablity, the successful prediction of surprising, previously unknown facts, and so forth. Put simply, folk theories are, while still theoretical in character (*qua* attempts to explain the world), not up to snuff by the criteria we now use to evaluate theories. Churchland is especially suspicious of folk theories for these very reasons, and historically they have been abandoned in favor of a radical new scientific framework. He expects folk (belief/desire) psychology will fare no better than folk astronomy or superstitious beliefs in the reality of witches.

As the central explanatory resources of folk psychology are intentionality ("aboutness") of the propositions that occur in natural language and thought, and our *attitudes* towards them (*believing* that proposition **P** is true, or *dreading* that it is), these essential "theoretical posits" are candidates for elimination. But without propositions themselves (occurring in either natural or a "mentalese" language), what of truth? Propositions (public language sentences, *et cetera*) are, on virtually any philosophical

account of language, the *bearers* of truth, the basic "units" which are capable of being true or false. Without there being propositions, nothing else in the world could be said to have the property of truth.

Even worse, if there are no propositions, there certainly are no *attitudes towards them*. Knowledge, as traditionally conceived by philosophers (from Plato onwards), crucially requires (at least) that there be at least one true proposition and at least one propositional attitude, namely, belief. How, then, am I able to take seriously Churchland's pleas that he is a scientific realist, in a robust sense? Scientific realism requires (at least) that we possess *knowledge* of some kind that the anti-realist proscribes. As we have seen, I have ardently maintained all along that we should erect a wall of strict separation between semantics and metaphysics/epistemology. But there is the troublesome fact that the following thesis is taken to be a platitude by an overwhelming majority of philosophers:

(Platitude **K**) Necessary (but probably insufficient) conditions for agent ***x***'s knowledge of proposition **P** include the following:

1. Agent ***x*** believes **P**.
2. Agent ***x*** has a sufficient degree or amount of evidence (warrant, justification, reasons, *et cetera*) for thinking that **P** obtains, that it describes what the world is like.

3. **P** is true, it obtains, it describes what the world is like, *et cetera*.

Platitude **K** requires that there be propositions, epistemic attitudes towards such propositions, and that propositions are truth bearers. Churchland rejects all three of these suppositions: propositions do not exist, therefore nothing can be said to be "true" (in a traditional sense, see below), and also therefore no one can *believe* anything (no one can have a meaningful attitude towards something that doesn't exist). The CRC seems to require that we *could* or actually do *know* which scientific theories or parts thereof are correct. Again, how can Churchland be a scientific realist?

The answer is that Churchland thinks that what we now call "knowledge," is just another word for a model or description of what the world is really like, one that "fits" or "matches" the structure of the world, or is a reliable guide to how it works, or (dare I say) "true." For now, "truth" is a placeholder for whatever accurate or correct description will turn out to be, as the forthcoming unarticulated replacement psychological theory has not yet been articulated. We don't yet have the (*neurological*, non-linguistic) tools to "say" what this will be. According to him, this advancement will occur not by lowering our standards but by "because we had *raised* our sights, in pursuit of some goal even *more* worthy than truth. I cannot now elucidate such goals [note: due to the aforementioned problem], but we should be sensible of their possible existence" (Churchland, 1985). (At least regarding his views on *mental representation*, knowledge in

a *private* sense, but probably not knowledge expressed in *public* language sentences, I am inclined to agree with Churchland on this point.)

Whether I agree with Churchland or not is immaterial for my present purposes. What (really) matters is that Churchland considers himself a scientific realist, and his emphasis on metaphysics and epistemology, over semantics, mirrors my own attitudes. So I am inclined to include his version of scientific realism as fitting under the "big tent" of the **CRC**, peculiar though it may be.

The second view I mentioned above, the more conservative view that sentential truth is a poor model of scientific representation, owes to Ronald Giere. Giere is a proponent of the "model-theoretic" or (poorly named) "semantic view" of theories, according to which (roughly), theories are sets of abstract models instead of the "received view" that theories are sets of axioms (ideally stated using only first-order language), supplemented with a "partial interpretation" into natural language, informing us how theoretical terms are to be understood using "correspondence rules" or "bridge principles," which link observational terms (which we use and understand unproblematically) to theoretical ones. The debate over these competing views of scientific representation is about as old as the debate over scientific realism, the dialectic is detailed and intricately subtle, and I cannot compare the views in the space I have here (although I will revisit this issue when I discuss structural realism). Put simply, the details of the debate deserve a chapter in its own right.

What the reader needs to know for our purposes is that according to the semantic view, scientific representation occurs not by way of (partially interpreted) theoretical sentences (stated in some first-order language) asserting universal generalizations about the behavior of the world. Instead, theories are (sets or collections of) abstract (usually mathematical) models. These mathematical models are "similar in respects and degrees" (Giere, 1988), or "partially isomorphic" to the "target system," the real-world system to be described. An example may help. Think, for instance, of the equation for the motion of a pendulum. The values of the parameters in this abstract equation model, to a certain degree, physical of the target system (determinate values of determinable properties of objects – in this case, the spatial position of the pendulum relative to a frame of reference, the force of gravity driving the oscillatory motion, the length of the wire, and so forth). Of course, in many ways, they don't – for instance, according to the equation as stated, the pendulum *would* oscillate forever and its maximum height would remain the same. In real life, air resistance dampens the gravitational force acting on the pendulum, eventually bringing it to rest. Of course, we can de-idealize by introducing a dampening factor into the equation, and thus have the theory (the mathematical model) represent the behavior target system more accurately. But perhaps there will always be idealizations involved. The ubiquitous use of fictional, ideal abstract entities (frictionless planes, perfectly elastic collisions, *et cetera*) is one

central motivation for Giere's preference of the model-theoretic account (another owes to doubts over "the bastard notion of approximate truth" (106)).

Hopefully, I have given the reader at least an adequate "sketch" of the debate. But adjudicating it does not matter here. What matters is that Giere, like Churchland, considers himself a scientific realist of a peculiar sort (given his agreement with van Fraassen on the importance of model construction in science, and the model-theoretic view of theories generally, his preferred term for his view is "constructive realism"). But he, like Churchland, is welcome under the CRC, even though he thinks truth doesn't enter into the very definition of realism. The same thread runs through both of their versions of scientific realism – it is "getting it right," faithfully representing what's in the world and why things happen, that matters. This may occur by way of our formulating true (or approximately true) sentences, it may occur by way of some non-linguistic mental representation "even more worthy than truth," or it may occur by similarity or partial isomorphism between abstract models and the real world. However representation happens, the scientific realist agrees, fundamentally, on one thing: that we do succeed in constructing representations that tell us what the world is really like. Epistemology – how we know and what we know about what the world is like, and how we can be so sure that we *do* have such knowledge, is the primary issues. Semantics, in my view, falls by the wayside, or it ought to. Scientific realists' historical *obsession* with the concept of truth, *especially as correspondence*, when this is a question best left for the

philosophy of language, is misguided. In so far as truth, especially correspondence truth, brings on otiose and unnecessary assumptions about the nature of representation and even ontology (e.g., real kinds, propositions, perhaps even a "spooky" relation of truth-making) it is best that we stay away from formulating the **CRC** as in any way involving the thesis that theories are true.

There one final matter I have not yet spoken to, the Verisimilitude Thesis. This is the idea that theories are not true *simpliciter* but merely approximately true. I am inclined to say that with my recommendation against semantic versions of scientific realism, approximate truth should drop out of the picture as well. But matters are (perhaps) a bit more complicated. I have consistently substituted phrases like "representational accuracy," "description of the way things really are," for the traditional realist favorite, "is true." But obviously, scientific theories do not give us a *fully accurate* or *completely trustworthy description*, in all respects, of how things really are. This is a fact we know: Newtonian mechanics is strictly false but works very well for most practical purposes, and perhaps "gravitational force" can be reduced to "space-time curvature" in some interesting way (smoothly or not). General Relativity and quantum mechanics are known to be inconsistent with each other, despite their fantastic records of explanatory and predictive success. So the Verisimilitude Thesis teaches us an important lesson. While science doesn't tell us what the world is really like up to every last little detail, even though it may represent the world as being one way when in fact

this is not so, theories are still our best guides to what exists, and why things happen. And, in some cases, our common-sense scientific realist believes that scientific theories are trustworthy representations of the world's structure *within circumscribed domains*, where we know where they fail to accurately represent (for instance, whatever theory supersedes General Relativity, it is very likely that the theory still correctly describes the large-scale and high-energy structure of space-time, where quantum effects are negligible).

Another point worth quick mention is something that has come up in the discussion of Giere's model-theoretic account of scientific representation, the importance of idealization in science. Most of the things scientists discuss (or physicists, at least) are outright fictions (a point hammered home by Cartwright (1983)). Consider the laundry list: two-body gravitational systems, frictionless planes, perfect vacuums, perfectly elastic solids, ideal gases, and so forth. Idealization is an essential feature of most of what goes on in science, *especially* at the level of deep, high-level theory. Not only is this a reason for doubting that theories are true, but also for stressing the role, and need, for approximation and de-idealization in science. Perhaps we should revise the **CRC** (as described so far), in light of this insight: Scientific theories are reliable guides to how the world works, under "specific conditions of idealization and approximation," conditions where we have a grasp on how to de-idealize the theory, in order to understand how it actually does accurately describe the world. We can do this because we understand

theories as being perfectly accurate description of an idealized world, one where the confounding causal influences that actually obtain in our world (borrowing from our list of examples; other planets in the solar system, friction, atmospheric pressure, deformations of material bodies, inter- and intra-molecular forces, *et cetera*) are not in play. This insight is well taken and should be incorporated into the **CRC**, perhaps as the one positive result to come from our discussion of how semantics relates to the realism debate. Theories are not perfect guides to what the world is like. They are at best imperfect, but merely roughly accurate guides to reality, and we occupy a very strong epistemic position regarding these theories *especially when* we understand where the theory fails to represent accurately, and most importantly, *why*.

Relating this back to my earlier discussion of Psillos (propositional, correspondence-truth based) account of verisimilitude, which remember was *contextual* in nature, in that it required theory-specific details to specify exactly when and where the theory should be seen as making a genuine knowledge claim, and when belief should be suspended, there is a considerable degree of concordance here. Indeed, Psillos suggests as much in his discussion of Giere's model-theoretic approach, but he also argues that his propositional view of theories and Giere's are essentially equivalent, substituting "true" for "…is partially isomorphic [or *similar*] to," and letting the collection of models just be the propositions of a theoretical system (274-5). I adjudicate this dispute here, as I don't think semantics is a substantive issue within this debate, as I

have argued. However, I do find it interesting and promising that whether Psillos is right or not about merely differing in notation rather than substance from Giere, each account of theoretical representation can offer a sketch of the details of how to accommodate the fact that while scientific theories supply us with deep knowledge with some things, they are continually evolving and incomplete, and we have a real need to know where to place our trust.

Conclusion

I have argued here throughout that epistemological considerations need to be placed front-and-center in the scientific realism debate. Specifically, I have identified what I believe to be the "core" commitment that defines scientific realism with the robust epistemological claim that we actually *do* know (and not just *could* know) some of the facts that our scientific theories purport to tell us. Consequently, *these* facts are the things we should be scientific realists about, if there are any such facts – if we should be scientific realists. I *have not* argued for scientific realism here – rather, I have tried to argue what a plausible but *minimal* scientific realism would look like; one that is as undemanding as possible with respect to one's other philosophical commitments. I argued for the priority of epistemology by way of comparison to the philosophical problem of skepticism, for, as I have tried to demonstrate, scientific anti-realism is at its core a selective form of philosophical skepticism, one that tries to separate the stock of human knowledge (or putative knowledge) into two categories: "ordinary" or "mundane" knowledge and "theoretical" knowledge. The *scientific* anti-realist grants, along with the *philosophical* ***realist***, that we do know some of what we take ourselves to know. But then the scientific anti-realist goes further, identifying distinctive reasons why we ought to think that science is somehow *unique* as a form of inquiry either in its methods or in the conclusions in draws, such that the former are unreliable, and hence the latter are likely to be systematically wrong or misleading. These reasons need not be

global (i.e., philosophical, topic-neutral), indeed they can be quite theory-specific, but they must be "in principle" considerations to think that there is some fundamental limit to what we can know. Scientific realism, on this view, is just the logical contrary of anti-realism: it affirms that at least some scientific methods and theories are reliable and should be counted among the things human beings know.

Scientific realists have wanted to add on to this "core commitment" several theses that go beyond this necessary requirement: that nature must be structured rigidly into a type-hierarchy of natural kinds, for instance, or that theories must be true in a correspondence sense (or even *true*, assuming that the vehicles of *theoretical* representation are linguistic, and hence the correctness or incorrectness of such representations consist in their being true). But these are not needed to block the anti-realist's core criticism of scientific knowledge – they do not add anything substantive to the core commitment – nor should they be considered part *of* scientific realism proper. As I argued in Chapter 5, their motivation stemmed in part from historically contingent factors that may have made sense in light of the milieu in which scientific realism first emerged, but are quite unnecessary *now*.

An epistemological **CRC**, and the rejection of substantive metaphysical and semantic positions, has a positive upshot for the scientific realism debate going forward. First, it is consistent with a wide variety of *other* philosophical commitments that probably count as disqualifying from the perspective of more robust, more demanding

forms of scientific realism. For instance, eliminative materialism is consistent with scientific realism as long as we don't insist on reading scientific realism as the idea that "scientific theories are (approximately) *true*," and as long as the eliminative materialist has some story to tell about how human beings form representations of their world that correctly or incorrectly portray the way things are in the world. This, in turn, leads to a second reason for endorsing the **CRC** – it bolsters (and at the same time) sharpens scientific realism's credentials and its resistance to criticism. One cannot undermine scientific realism, for instance, by attacking a certain view of how theoretical terms get their reference, or by attacking the hypothesis of natural kinds, at least not without some additional argument (whose premises the scientific realist may not be willing to grant) connecting these ideas to the **CRC**. At the same time, identifying scientific realism with the **CRC** doesn't make scientific realism unassailable as a philosophical thesis, true come what may – and therefore, the worry might go, without content. As I said, the **CRC** *sharpens* scientific realism by squarely focusing the debate where it should be: on the question on whether and how we can take ourselves to have certain kinds of scientific knowledge. That is most definitely a falsifiable position, and one that is inconsistent with most extant versions of scientific anti-realism.

References

Alspector-Kelly, M. (2001). Should the Empiricist Be a Constructive Empiricist? *Philosophy of Science* , 68(4):413-431.

Aronson, J. A., Harre, R., & Way, E. C. (1995). *Realism Rescued: How Scientific Progress is Possible.* Chicago, IL: Open Court Press.

Boyd, R. (1999). Homeostasis, Species and Higher Taxa. In R. Wilson (Ed.), *Species: Interdisciplinary Essays* (pp. 141-186). Cambridge: MIT Press.

Boyd, R. (1983). On the Current Status of the Issue of Scientific Realism. *Erkenntnis* , 19:45-90.

Callender, C., & Magnus, P. (2004). Realists Ennui and the Base Rate Fallacy. *Philosphy of Science* , 71:320-338.

Carnap, R. (1956). Empiricism, Semantics and Ontology. In R. Carnap, *Meaning and Necessity: A Study in Semantic Modal Logic* (pp. 20-40). Chicago, IL: University of Chicago Press.

Cartwright, N. (1983). *How the Laws of Physics Lie.* Oxford: Oxford University Press.

Chakravarrty, A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable.* Cambridge: Cambridge University Press.

Churchland, P. (1981). Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy* , 78:67-90.

Churchland, P. (1985). The Ontological Status of Observables: In Praise of the Superempirical Virtues. In Churchland & Hooker (Eds.), *Images of Science* (pp. 35-47). Chicago, IL: University of Chicago Press.

Dupre, J. (1993). *The Disorder of Things.* Cambridge: Harvard University Press.

Einheuser, I. (2011). Toward a Conceptualist Solution of the Grounding Problem. *Nous* , 45(2):300-314.

Ellis, B. (2001). *Scientific Essentialism (Cambridge Studies in Philosophy).* Cambridge: Cambridge University Press.

Ereshefky, M. (1998). Species Pluralism and Anti-Realism. *Philosophy of Science* , 65:103-120.

Field, H. (1970). Theory Change and the Indeterminancy of Reference. *Journal of Philosophy* , 14:462-481.

Fine, A. (1986). *The Shaky Game: Einstein Realism and the Quantum Theory.* Chicago, IL: University of Chicago Press.

Fodor, J. (1974). Special Sciences: Or the Disunity of Science as a Working Hypothesis. *Synthese* , 28:97-115.

Fraassen, B. v. (1980). *The Scientific Image.* Oxford: Oxford University Press.

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

Hacking, I. (1999). *The Social Construction of What?* Cambridge: Harvard University Press.

Hardin, C., & Rosenberg, A. (1982). In Defence of Convergent Realism. *Philosophy of Science* , 49:604-615.

Hennig, W. (1966). *Phylogenetic Systematics.* Urbana, IL: University of Illinois Press.

Hume, D., & Beauchamp, T. (1999). *An Enquiry Concerning Human Understanding.* Oxford: Oxford University Press.

Kitcher, P. (2001). Real Realism: The Galilean Strategy. *The Philosophical Review* , 110:151-197.

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

Kuhn, T. (1970). *The Structure of Scientific Revolutions (2nd Ed.).* Chicago, IL: University of Chicago Press.

Kukla, A. (2000). *Social Contructivism and the Philosophy of Science.* London: Routledge.

LaPorte, J. (2004). *Natural Kinds and Conceptual Change.* Cambridge: Cambridge University Press.

Laudan, L. (1981). A Confutation of Convergent Realism. *Philosophy of Science* , 48:19-48.

Laudan, L., & Leplin, J. (1991). Empirical Equivalence and Underdetermination. *Journal of Philosophy* , 88:449-472.

Lee, S. (1995). Self-Starvation in Context: Towards a Culturally Sensitive Understanding of Anorexia Nervosa. *Social Science and Medicine* , 41(1):25-36.

Leeds, S. (2007). Correspondence Truth and Scientific Realism. *Synthese* , 159:1-21.

Leplin, J. (1997). *A Novel Defence of Scientific Realism.* Oxford: Oxford University Press.

Lewis, D. (1983). New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61:343-377.

Maddy, P. (2007). *Second Philosophy: A Naturalistic Method.* Oxford: Oxford University Press.

Mayr, E. (1942). *Systematics and the Origin of the Species.* New York, NY: Columbia University Press.

Musgrave, A. (1989). NOA'S Ark - Fine for Realism. *The Philosophical Quarterly* , 39:383-398.

Papineau, D. (2010). Realism, Ramsey Sentences and the Pessimistic Meta-Induction. *Studies in History and Philosophy of Science* , 41:375-385.

Plato. *Phaedrus 265d-266a.*

Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth.* London: Routledge.

Putnam, H. (1981). *Reason, Truth and History.* Cambridge: Cambridge University Press.

Schaffer, J. (2007). Causation and Laws of Nature: Reductionism. In Hawthorne, Sider, & Zimmerman (Eds.), *Contemporary Debates in Metaphysics* (pp. 82-107). Malden, MA: Blackwell Press.

Sellars, W. (1962). Philosophy and the Scientific Image of Man. In R. Colodny (Ed.), *Frontiers of Science and Philosophy* (pp. 35-38). Pittsburgh: University of Pittsburgh Press.

Sklar, L. (2000). *Theory and Truth.* Oxford: Oxford University Press.

Stanford, K. (2006). *Exceeding our Grasp: Science, History, and the Problem of Unconceived Alternatives.* Oxford: Oxford University Press.

Turner, D. (2007). *Making Prehistory: Historical Science and the Scientific Realism Debate.* Cambridge: Cambridge University Press.

Biography

Nathan Morton is 29 years of age and currently resides near Philadelphia PA, with his loving family: his wife, Samantha, his dog, Beauregard, and his year-old daughter, Morah Dail Morton. He grew up in this region and attended the University of Pennsylvania (B.A., Philosophy) before entering Duke University's Philosophy Ph.D. program. Nathan's primary research interests are the general philosophy of science, epistemology, and metaphysics, especially where these intersect each other. Upon reading Thomas Kuhn's *The Structure of Scientific Revolutions* over summer break in college, Nathan became intensively interested in the scientific realism debate, especially from a historical perspective. This encounter motivated him to change majors to philosophy, and Nathan has been thinking and writing about scientific realism ever since.