9 Scientific Fictions as Rules of Inference

Mauricio Suárez

1. FICTIONALISM IN THE PHILOSOPHY OF SCIENCE

Hans Vaihinger’s philosophy of “as if” introduced fictionalism into philosophical discussions in the early years of the 20th century. Vaihinger’s thesis was radical: A knowledge worth pursuing is thoroughly infused by fictional assumptions. Vaihinger distinguished carefully fictions from hypotheses, and considered most of science and mathematics to engage shamelessly in the production, dissemination, and application of both. Hypotheses are directly verifiable by experience and their truth is tentatively granted. Fictions are, for Vaihinger, accounts of the world and its systems that not only are plainly and openly false, but knowingly so, yet remain indispensable in theorizing—in science and elsewhere. However the fictions employed in scientific reasoning are not of the same kind as those that appear in other areas of human endeavor. Vaihinger distinguished scientific fictions from other kinds of fictions (such as poetic, mythical, or religious fictions), and he understood the difference to be one of function. Virtuous fictions play a role in a particular kind of practical rationality in scientific theorizing, a kind of “means–end” rationality at the theoretical level. In Vaihinger’s terminology, they are expedient. These are the fictions that figure in the scientific enterprise, and among the most prominent throughout the history of science Vaihinger identified forces, electromagnetic “lines” of force, the atom, and the mathematical infinity, as well as some of the main constructs of differential analysis such as infinitesimal, point, line, surface, and space.

Vaihinger’s work is unfortunately not sufficiently well known today, but he should appear to philosophers of science as an extremely contemporary figure. A recent brief paper by Arthur Fine brings Vaihinger back to the philosophical fore. Fine notes that “Vaihinger’s fictionalism and his “as if” are an effort to make us aware of the central role of model building, simulation, and related constructive techniques, in our various scientific practices and activities” (Fine, 1993, p. 16). There has been in the last decade or two a remarkable resurgence of interest in the topic of modeling, which emphasizes the essential role played by idealizations, contradictions, abstractions, and simulations in the practice of scientific modeling. Vaihinger’s work fits
right in: His concern to emphasize the use of false or contradictory assumptions in building models of systems and their workings is of a piece with this whole body of literature. So too is his concern to appreciate the pragmatic virtues that these assumptions might bring to scientific reasoning. For instance, Vaihinger also distinguished between fictions (which involve internal contradictions, or inconsistencies) and what he called semi-fictions (which involve contradictions with experience).⁶ We would nowadays refer to the former as contradictions, and claim that they are logically false, while calling the latter idealizations, which we would claim are empirically false. This distinction will be appealed to freely in the text, and instances of both types of fiction will be identified. But I will not be concerned very much with the fact that fictions in science tend to entail falsehood. Although Vaihinger emphasized the falsehood of scientific fictions, his main concern was with their cognitive function.⁷ The main interest throughout this chapter is in shedding some light on the cognitive function of scientific fictions, regardless of their truth value. In particular I urge that expediency in inference is the main defining function of a scientific fiction. This chapter is a first attempt at an elaboration and defense of this inferential expediency of scientific fictions. It can be seen as part of a larger argument for taking the cognitive value of a scientific fiction to lie entirely in its function in inquiry, and to be fully independent of its truth value.

The modeling scholarship of the last decade or so has made a strong case for a version of Vaihinger’s main thesis: The use of fictions is as ubiquitous in scientific narratives and practice as in any other human endeavor, including literature and art; and scientists have demonstrated throughout history a capacity to create, develop, and use fictions for their own purposes that compares with that of any writers or artists. This is why this assumption will not be defended here, but will rather be taken as a matter of fact, and a starting point. The aim is instead to explore a philosophical issue that is key to the success of the philosophy of modeling movement, and to Vaihinger’s conception in particular. This is the fundamental distinction between what Vaihinger identified as scientific fictions and other kinds of fictions.⁸ It is this distinction that lies at the heart of the thesis that fictions are of particular use to science, or that science employs them in a particularly useful way. Without an elaboration and defense of the expediency that characterizes scientific fictions the main thesis, however true, would bear little content—and it would shed little insight on scientific practice.

2. FICTIONS IN THE HISTORY OF SCIENCE: TWO EXAMPLES

In order to characterize scientific fictions we may begin with some illustrious examples of productive use of fictions in the history of science, the ether theories and the models of the atom at the turn of the century. They are interesting in complementary ways. In the ether case the use
in mathematical and experimental inference of what we nowadays take
to be a fictitious entity was extremely productive and long-lasting—
although convictions as to its reality differed and wavered considerably.
The case of atomic models illustrates another possibility: a model of a
putatively real entity that was never taken very seriously—the model was
never assumed to be other than a fictional description of the entity—and
was consequently short-lived, yet turned out to be extraordinarily useful
as a heuristics for developing new, more detailed and powerful models.
Therefore these cases illustrate the difference between representations of
fictional entities and fictive representations of real entities; they also illus-
trate the difference between long-lasting formal tools and short-lived con-
crete models. Yet Vaihinger’s thesis holds in both cases, because fictions
are involved in an essential way, one way or another. The description that
follows will emphasize some features that in both cases seem relevant to
the distinction between scientific and nonscientific fictions. (Ultimately
the philosophical discussion will focus on a further example from quan-
tum theory.)

Throughout the 19th century the ether was taken to be the putative
carrier of the light waves, and stellar aberration phenomena had estab-
lished that the earth must have been in motion with respect to it. The
most sophisticated mathematical theories of the ether were developed in
Britain in the wake of Maxwell’s theory of electromagnetism. Maxwell
developed his theory in the years 1856 to 1873, the year of publication
of his Treatise on Electricity and Magnetism.9 The theory went roughly
through three phases corresponding to the publication of “On Far-
aday’s Lines of Force” (1856), “On Physical Lines of Force” (1861), and
finally “A Dynamical Theory of the Electromagnetic Field” (1865) and
the Treatise. In each step Maxwell’s theoretical model acquires a higher
degree of abstraction from the mechanical model of a luminiferous ether.
The central concept throughout is that of field energy: the only concept
toward which Maxwell shows an increasingly firmer ontological com-
mitment over the years. By contrast, Maxwell’s reluctance to accept the
ontological implication of the existence of the mechanical ether is strong
and explicit from the beginning. As Morrison has pointed out, in drawing
the analogies in “On Faraday’s Lines of Force” between electrostatics,
current electricity and magnetism with the motion of an incompressible
fluid, this fluid “was not even considered [by Maxwell] a hypothetical
entity—it was purely fictional” (Morrison, 2001, p. 65). And later on as
he developed a clearer view of what have come to be known as Maxwell’s
equations, Maxwell remained resolutely skeptical regarding the existence
of the ether. The mechanical models of the ether were gradually stripped
of ontological content as the dynamical equations were developed, yet
these mechanical models remained indispensable to scientific theorizing,
according to Maxwell, “as heuristic devices, or at best, descriptions of
what nature might be like” (Morrison, 2001). In other words, the ether
remained for Maxwell a fiction—with an essentially heuristic function in
the development of increasingly sophisticated mathematical models that
captured the empirical phenomena. This was its scientific virtue—to be
a guide in the construction of more detailed, explanatory or predictive
models of the phenomena.10

The Maxwellian tradition was continued among theoreticians in Cam-
bridge, mainly through the work of Charles Niven and Joseph Larmor.
Andrew Warwick has written a scholarly and detailed history of the
Cambridge couching in electromagnetic theory from the very first lect-
ures that Maxwell himself gave there on publication of his Treatise until
the final demise of ether-based theories in favor of Einstein’s theory in
the 1920s (Warwick, 2003). What is remarkable about this history is the
role that the ether played as basis for the application of electromagnetic
theory to all kinds of practical and experimental problems. Although the
Cambridge Maxwellians professed a belief in the ether (and thus, unlike
Maxwell himself, took the ether to be a hypothesis rather than a fic-
tion in Vaihinger’s terms), they nonetheless employed the ether as a men-
tal construction that allowed them to apply electromagnetic theory in a
much simpler and straightforward way. The critical difference is that the
ether theories allow for electromagnetic effects to be present all over as
due almost entirely to the flow of energy “in the ether,” while action-at-
a-distance theories supposed that electromagnetic effects only occurred
within conducting and dielectric material, not at all in empty space.
As Warwick puts it, “the solution of problems using the field-theoretic
approach therefore required very careful consideration of the electromag-
netic action and boundary conditions applicable at or near the surface of
conductors” (Warwick, 2003, p. 329). Confidence in the applicability of
the equations is required for carrying out these calculations, but a belief
in the existence of the ether is not strictly required. In fact, convictions as
to the existence of the ether waivered a great deal, and differed between
its different proponents.11

In other words, the fiction—or hypothesis—of the ether has for both
Maxwell and Larmor a powerful positive heuristic for further research,
giving guidance on how to generate more detailed models of the phenom-
ena, allowing quick and effective inference of experimental and practical
results that can then be readily checked against experience. It provides
an excellent tool for the calculation of effects and the regimentation of
knowledge. It also provides what Mary Hesse aptly called neutral analog-
ies—unproved similarities between the ether models and real phenom-
ena that call for exploration and investigation.12 But the reality of the
ether itself is not required for inference, taxonomy, or analogy. The ether
provides a mental model of great expediency and ease of calculation, but
it can well remain “a mere figment of the mind, not a fact of nature”
(Maxwell, on referring to disturbances of the ether, as quoted in Mor-
ris, 2000, p. 96).
The second example is even more startling in the fictitious nature of the description employed: J. J. Thomson’s *plum pudding* model of the atom. Although the entity described is by no means considered a fiction by our present lights, the description of the entity given in the model in question certainly is fictional and in ways interesting to our present purposes. The idea that electric conductivity phenomena may be underpinned by the fundamentally asymmetric nature of the distribution of positive and negative charges in atoms actually goes back to Lord Kelvin, who proposed it in his President’s Address to the Royal Society in 1893. In fact, Kelvin had already proposed something very similar to a plum-pudding model for the atom as part of his defense that Crookes’s experiments on cathode rays established that they were negatively charged particles. Thus Thomson did not “make up” the plum pudding model, but was following in Kelvin’s footsteps in proposing and developing it in the years between 1897 and 1910. The model was part and parcel of the strategy followed by both Kelvin and Thomson in trying to prove the particulate nature of cathode rays, and the fundamental electrical asymmetry of conductivity phenomena.

According to the plum pudding model the atom is a roughly spherical *sponge* formed by evenly distributed positive charge in which minute negatively charged particles (“electrons”) are inserted, like raisins in a traditional British Christmas cake. The model had the great advantage of explaining ionization phenomena, whereby negative charged particles are bounced off atoms by collisions with other atoms, creating electric currents. It also explained the production of cathode rays and sustained Kelvin’s and Thomson’s favored interpretation of them as negatively charged particles. But most importantly, the plum pudding model was a powerful heuristics in the development and application of Thomson’s “working hypothesis” between 1897 and 1910 for the experimental inquiry into electrical phenomena in gases, namely, “that the negatively charged corpuscle is universal and fundamental, ionization results from the dissociation of a corpuscle from an atom, and electrical currents in gases at low pressures consist primarily of the migration of corpuscles” (Smith, 2004, p. 23).

As is well known, the plum-pudding model was refuted by the experiments performed in 1909 by one of Ernst Rutherford’s collaborators in Manchester, Hans Geiger, together with a student, Ernst Marsden. In these experiments thin foils of gold were bombarded with alpha particles (essentially a couple of protons and neutrons bound together as in helium nuclei) and a significant proportion of recoil was observed. (On the order of 1 in 20,000 alpha particles was deflected by an average angle of 90 degrees.) The effect is sometimes known as Rutherford scattering, and was effectively employed by Rutherford in order to reject the plum pudding model because the electrons in the atom are too small and light to produce the recoil effect, while the diffuse “sponge” would be porous to the massive and energetic alpha particles. Hence Rutherford advanced his hypothesis of
the existence of a massive and discrete nucleus at the center of empty space, orbited by minute electrons, in order to explain the scattering effect with the experimentally observed probability.

It would be uncharitable, however, to charge Thomson with the belief that his plum pudding model was a complete and true description of the atom (see Smith, 2004, particularly p. 25). Instead the model served three different heuristic functions: (i) it allowed Kelvin and Thomson to argue for the particulate as opposed to wave-like nature of cathode rays (electrons), a fact that Thomson’s 1897–1899 papers were instrumental in establishing universally; (ii) it justified and provided motivation for Thomson’s “working hypothesis” regarding conductivity in gases during the 1897–1910 years, which in turn led to many important experimental results and conclusions—including among others the establishment of the existence of electrons themselves; and (iii) it helped Rutherford target the experimental evidence available toward the missing or defective assumptions in the previous models, thus enabling him to discover the existence of the nucleus. (As is well known, Rutherford’s “planetary” model was in turn superseded by Bohr’s early quantum model [1913], which hypothesized spontaneous quantum transitions between electron orbitals, and hence a nonclassical structure of quantum energy levels within the atom.)

It is hard to overestimate the importance of the plum pudding model in helping focus inference making in all these three different areas. The historical facts rather point out that these three crucial and critical developments and discoveries in the history of modern physics could hardly have come about without the expediency in reasoning provided by the plum pudding model, no matter how fictitious. The model served an essential pragmatic purpose in generating quick and expedient inference at the theoretical level, and then in turn from the theoretical to the experimental level. It articulated a space of reasons, a background of assumptions against which the participants in the debates could sustain their arguments for and against these three hypotheses. As Thomson himself put it: “My object has been to show how the various phenomena exhibited when electricity passes through gases can be coordinated by this conception.”

3. FICTIONS IN CONTEMPORARY PHYSICS

The use of fictions is not confined to past or failed science. Let us now consider a couple of typical instances of fictions within contemporary successful science. First, an example from astrophysics is briefly discussed: the Vaihingerian semi-fictions involved in models of stellar structure. Then we turn to a striking case of Vaihingerian full fiction: the models of the theory of quantum measurement. Once again these case studies emphasize the inferential function of the fictional assumptions involved in both models.
3.1. The Semi-Fiction of a Star

Models of stellar structure in astrophysics provide a description of the inner workings of a star; in particular the fuel burning processes (nuclear fusion) that turn hydrogen into helium and generate the star radiation, while accounting for the star life cycle and evolution. These models match up the observational quantities of a star, which mainly pertain to the properties of its photosphere, that is, the outermost layer of the star. The observable quantities of stars include: (i) its luminosity (energy radiated per unit time, which depends upon apparent brightness and distance), (ii) its surface temperature (the temperature of the photosphere), (iii) the photosphere’s chemical composition, and in rare cases (binary stars) the mass of the star. The models allow us to infer conclusions regarding the internal workings of a star on the basis of these observational quantities. Their cognitive value depends upon the reach, power, and ease of calculation of such inferences.

These models make at least four assumptions that are widely assumed to contradict either the physics of matter and radiation, the physical conditions of the interstellar medium, or both. Hence the models are knowingly strictly speaking false of real stars. I won’t discuss them here in detail, but will just briefly discuss these four assumptions (see, e.g., Prialnik, 2000, pp. 6–8. Also Tayler, 1970). First, models of stellar structure assume that a star is an isolated bubble of gas; that is, they assume that a star is a physically closed system—no external forces (gravitational or electromagnetic) intervene. The only forces that can affect the internal structure of a star are consequently assumed to be the star’s internal gravitational forces due to the rotational movement of the gas, and the forces that arise out of the nuclear burning inside the star. As far as external gravitational forces go, the assumption of isolation is not entirely unrealistic: A star is a dense concentration of hydrogen in the interstellar medium, and the nearest such concentration can be on average as far as 4.3 light years away (this is roughly the distance between the sun and Alpha Centauri, the star nearest to our sun). So external gravitational forces are bound to be tiny, and therefore negligible for all practical purposes. But the assumption of isolation is physically quite incorrect: The interstellar medium is not empty space but is replete with irregularly dense (although on average much lighter) gas, mainly hydrogen, which in the surroundings of the star is gravitationally attracted into the star, while radiatively repelled. In other words, isolation suggests that the physical boundaries of a star are sharp, when as a matter of fact they are rather imprecise.

Second, stellar structure models assume that all stars possess identical chemical composition, namely, the sun’s: 70% hydrogen and 30% helium. The presence of heavier elements has been confirmed in every star, but it is essentially irrelevant to the structure of the star in most hydrogen-burning models, so it is entirely neglected. In these models, a small difference in the proportion of hydrogen and helium in the initial composition of a star can
have an important effect in its evolutionary life (it can particularly affect its luminosity and lifetime), but because the models have other parameters to adjust for these observable quantities, these differences are essentially ignored, and a blanket assumption is made.

Third, it is assumed that the shape of the star is spherically symmetrical throughout its life. Yet both internal rotational forces and magnetic forces are well-known causes of departures from spherical symmetry. Hence this assumption rules out such internal forces, which as a matter of fact are known to be large. Finally, a star is assumed to permanently stay in a state of thermal equilibrium, whereby the temperature of the gas is identical to the temperature of the radiation emitted (see Prialnik, 2000, Chapters 3 and 5). The implication of this assumption is that radiation provides the only form of energy transfer within a star. It also follows that the energy spectrum of a star is a black-body spectrum. Yet, massive convective fluxes are known to occur in periods of heavy hydrogen burning within the star. So the assumption is that such convective forces have no impact on the temperature distribution inside the star. Because the inside of a star is entirely theoretical, the only justification for this assumption is the model’s ability to correctly match the observable quantities of the photosphere.

These assumptions are not discussed any further here. I merely emphasize that they are all known to be empirically false, or at any rate unproven, yet in combination they afford a huge improvement in the expediency of the inferences that can be drawn from the models to the observable quantities, and between such quantities Their effect on the accuracy of predictions from the model cancels out, and is therefore negligible in comparison. In other words, the stars of contemporary astrophysics’ models are Vaihingerian semi-fictions whose justification lies entirely in the great ease and expediency in inference making that they generate.17

3.2. The Full Fiction of Quantum Measurement

An example from present-day theoretical science is next discussed that exhibits the characteristic expediency of fictions in a prominent and illustrative manner. The model that quantum theory provides for measurement interactions is intended as a representation of a physical process—an image of a physical process provided by a highly formal and mathematical model. It might seem surprising that I am characterizing it as a fiction, because the formal machinery of quantum theory is so solidly entrenched among practicing physicists. And yet, its fictional character has actually been proved as a mathematical theorem—known as the insolubility proof of the measurement problem. First the basic assumptions underlying the model are discussed (sec. 3.2.1), and then reasons are provided for considering it a semi-fiction in Vaihinger’s sense (sec. 3.2.2), a maximal semi-fiction, and a fully fledged fiction (sec. 3.2.3). Sections 3.2.1 and 3.2.2 are technically a little demanding (even though the most demanding technicalities have been
confined to an appendix), and the uninformed reader may well want to skip them without loss.

3.2.1. The Quantum Theoretical Model of Measurement

Quantum theory provides an abstract mathematical model of physical interaction between a microscopic quantum object and a macroscopic device designed to test the state of the object. The theory, as first formulated by von Neumann (1932), ascribes a quantum state to the measuring device, and treats the interaction as a quantum interaction, that is, one that obeys the Schrödinger equation.

Suppose the initial state of the system is \( W_o = \sum_n p_n P[\nu_n] \), where each \( \nu_n \) may be expressed as a linear combination of eigenstates of the observable \( O \) of the system that we are interested in (i.e., \( \nu_n = \sum c_n | \phi_n \rangle \)); and that the initial state of the measuring device is \( W_a = \sum_n w_n P[\gamma_m] \). Throughout the chapter I refer to the observable represented by the operator \( I \otimes \Lambda \), as well as that represented by \( \Lambda \), as the pointer position observable. The eigenvalues of this observable are the set \( \{ \mu_{\alpha} \} \). As the interaction between the object system and the measuring device is governed by the Schrödinger equation, there must exist a unitary operator \( U \) that takes the initial state of the composite system (object system + measuring device) into its final state at the completion of the interaction, as follows: \( W_o \otimes W_a \rightarrow U (W_o \otimes W_a) U^\dagger \). (For further details of the interaction formalism, see Appendix 1.)

3.2.2. The Problem of Measurement

The intuition behind the so-called “problem of measurement” is easy enough to state. Take a system in an arbitrary superposition \( \nu_n = \sum c_n | \phi_n \rangle \). Then, due to the linearity of the Schrödinger equation, at the conclusion of an ideal measurement interaction with a measurement apparatus in any pure state, the composite (system + device) will be in a superposition of eigenstates of the pointer position observable. And according to the so-called eigenstate–eigenvalue link (e/e link), the pointer position observable cannot have a value in this state, because it is not in an eigenstate of the relevant observable. But surely quantum measurements do have some values—that is, they have some value or other. Hence the quantum theory of measurement fails to describe real quantum measurements, and the model expresses a fiction—to be exact, a semi-fiction in the terminology of Vaihinger, because it contradicts empirical reality.

3.2.3. The Model is a Maximal Semi-Fiction

Let us refer to a model as a maximal semi-fiction in Vaihinger’s sense if every assumption in the model can be shown to be false. Then the quantum theoretical model of measurements is a Vaihingerian maximal semi-fiction.
All the assumptions required for a quantum theoretical model of measurements are strictly speaking false. In addition to the eigenstate–eigenvalue link and the assumption that the Schrödinger equation is the full dynamical description of events, there are two formal conditions that are required to derive the problem of measurement. I have elsewhere referred to them as the transfer of probability condition or (TPC), and the occurrence of outcomes condition, or (OOC). (They are both described formally in Appendix 2.) Informally, (TPC) states that the probability distribution over the eigenvalues of the initial object system should be reproduced as the probability distribution of the pointer position observable. (OOC) by contrast states that the final state of the composite is a mixture over states in which the pointer position observable takes a particular value or other with probability one, and is often thought to be inspired by the eigenstate–eigenvalue link.

(TPC) is strictly speaking false on at least two counts. First, it assumes that whether interactions are measurements is an all-or-nothing affair that does not depend on the actual initial state of the system to be measured at a particular time, but on all the possible states that the object may have had in accordance with the theory. This is hardly satisfied by any real measurement we know. For instance, in setting up a localization measurement of the position of an electron in the laboratory, we do not assume that the device should be able to discern a position outside the laboratory walls, even if it is theoretically possible that the particle’s position be infinitely far away from us. All real measurement devices are built in accordance to similar assumptions about the physically possible, as opposed to merely theoretically possible, states of the object system, on account of the particular conditions at hand. So real measurement devices do not strictly speaking ever fulfill (TPC).

Second, (TPC) appears to require measurements to be ideal in the technical sense of correlating one-to-one the initial states of the object system with states of the composite at the end of the interaction. However, many real measurements are not ideal in this sense. Most measurement apparatuses make mistakes, and no matter how much we may try to fine-tune our interaction Hamiltonian, we are likely in reality to depart from perfect correlation. So again, in most cases of real measurements (TPC) will not apply.

Let us now turn to (OOC). This is also strictly speaking false, because it assumes that the measuring device can only “point” to the eigenvalue of the pointer position observable that has probability one in the final state that results at the end of the interaction. But we can see that this assumption is false in most measurement interactions with quantum objects in mixed states where outcomes are produced, and pointers “point” in spite of the probabilistic nature of the transitions.

Finally, the application of the Schrödinger equation in this setup is also strictly speaking false because it implies the assumption that all quantum systems, not only composite systems involving measuring devices, are
closed systems. It assumes that the quantum Hamiltonian can transform pure states into pure states, or mixtures into mixtures, but never a pure state into a mixture or vice versa. In reality, all systems are open and subject to a degree of state shift due to interaction with the environment and background noise—this phenomenon is known as decoherence.

3.2.4. The Model Is a Full Fiction

But in fact the insolubility proof of the quantum measurement problem shows that these three premises (together with a fourth premise named RUE, which I have ignored here) are inconsistent under the standard interpretation of quantum observables, the so-called eigenstate-eigenvalue link, or e/e link.\(^1\) Hence the model is not only empirically false, and maximally so, but it turns out to be necessarily false, because it is internally incoherent. In Vaihinger’s terminology, the quantum theoretical model of measurements is a fully fledged scientific fiction.

3.2.5. The Model Is a Scientific Fiction

The main feature of the model, shared with most theoretical models, is its inferential capacity. Once we understand how the model works we are in a position to draw inferences regarding the typical final state of a composite (object + measuring device) after the interaction. From this state we can further infer the possible values of the pointer position observable of the measuring device at the conclusion of the interaction (as long as the final state has a particular form that allows us to do just that). And from this inference we finally come to understand that there is an inconsistency in the account, which in turn leads us to consider which among the assumptions might be false.

Similarly, we have a large array of applications of measurement theory to different cases of physical interactions at the quantum level, each of them relaxing different assumptions of the model. For example, by relaxing the (e/e link) we have models within the modal interpretation of quantum mechanics; by relaxing the assumption of Schrödinger evolution of the composite we derive stochastic reduction and quantum state diffusion models; by relaxing the (OOC) assumption we obtain statistical interpretations of the measurement interaction, and hidden variable models. Finally, relaxing (TPC) allows us to encompass and develop models of highly non-ideal interactions, such as destructive measurements.\(^1\)

As in the previous example, the fictitious entity or process—the ether described by ether theories, the atom as plum-pudding, the internal structure of stars, measurement interactions as described by the quantum model—has in the appropriate formal model the capacity to generate various inferences to experimental or practical results, some of which can then be tested against experience. It is not the fiction itself that is in the first
instance under experimental test, but the results of inferences that the fiction licences in appropriate formal frameworks, or in conjunction with further assumptions and background theory. In all these cases the inferences would either be arbitrary or impossible to even draw without the supposition of the fiction in the first instance. The fictitious entities or processes articulate a framework for quick and expedient inference-making that would be either impossible or arbitrary otherwise.

4. A SUBSTITUTIVE FUNCTIONAL ACCOUNT OF SCIENTIFIC FICTIONS

On the account defended here the hallmark of scientific fiction is expediency in inference. Note that this is not equivalent to the claim that the fictions employed in art and literature, unlike those from science, do not possess a capacity to generate useful (imaginative, pleasant, interesting) inferences. It is obvious that literary and artistic fiction can and must serve that purpose too. But expediency is not generally considered a virtue. To the contrary, it is often derided for good literary fiction. Nor is it ever required in such cases that the conclusion of our inferences be at least in principle testable against experience. There is thus a double norm that rules the use of fictions in science in comparison with nonscientific fictions. First, scientific fictions will be judged by their capacity to allow expedient inference-making. Second, at least some of the conclusions arrived at by means of such inferences will be taken to be empirically testable at least in principle. Most fictional assumptions in science will possess both functional virtues. By contrast, the fictions of art and literature can at best share in the first kind of virtue, but they need not even do so in order to fulfill their proper aesthetic function.

So how do these “fictitious” representations in science work? There is by now a large literature on the topic of representation, particularly as applied to artistic or pictorial representation. A comparison with such theories will be illustrative. The use of fictions in science seems to share some elements in common with Gombrich’s substitutive account of fictions in art, and his account of pictorial representation more generally. On this account, the function of pictures is to cognitively replace the objects that they depict—the replacement has a cognitive function in the sense that it allows surrogate reasoning. In reasoning about the properties of the picture, we may be able to infer properties of the object depicted. Gombrich employs the analogy of a hobbyhorse, which children play with in a similar substitutive fashion. In this example, the replacement is not merely cognitive, but also of a practical nature (often missing in the painting and in the scientific case), so actions can be performed around the hobbyhorse that would be performed around a real horse. In these activities children can sometimes lose track of the fictional nature of the entity—in fact there
is a sense in which for it to perform its function correctly, it is essential that the fictional nature of the entity be in some ways suppressed. Although in Gombrich’s substitutive account of representation we are not required to gain the (false) belief that the fictitious entity itself exists, it seems that we are required to display at least some attitudes toward the fiction that we would display toward the real entity. In the hobbyhorse case this “pretence” attitudinal set includes practical action; in the painting case it merely includes cognitive factors.

Something very similar operates in the case of scientific fictional representation. What we do—what scientists do—when employing a fiction for scientific purposes—what Maxwell does in employing the ether, Thomson in employing the plum pudding model, or quantum physicists in applying quantum measurement theory—is to substitute the real process (with all its complications, disturbing factors, and exogenous causes) with a simpler, streamlined, and coarser-grained fictional account (one where all the factors and causes are fully and carefully identified), which is more expedient in the practice of reasoning and inference-making. We then go on to investigate the properties of the substitute rather than the real process. It will be typical for many scientists engaged in that process to lose track of the fact that the entities invoked are fictitious. This is as it should be on a substitutive account, because the representational success of the model depends on its success in generating the same set of cognitive attitudes toward it that scientists would exhibit toward the entity modeled. The attitudes of Larmor and Trouton toward Maxwellian ether-based electromagnetism just reflect this: Although there was never any experimental confirmation of the ether (and Larmor himself thought such confirmation was made impossible by the theory itself—i.e., by the ether-fiction itself), a realist attitude toward the ether assumption was a powerful heuristic in applying the theory to the most diverse phenomena. But although a realist cognitive attitude will be typical, a belief in the fiction is not necessary in order to operate the model successfully—we already saw that Maxwell himself displayed a much more cautious attitude toward the ontological import of the ether.

And mutatis mutandis in the quantum measurement case: To model quantum measurements in this way involves the cognitive attitude toward the model that we would adopt if confronted with real measurement processes. Thus it becomes possible to draw inferences regarding the real outcomes of real measurements from the model. But a belief that real quantum measurement interactions fully obey the formal assumptions is not required for such a cognitive attitude, which can be displayed simultaneously with a belief in the fictitious character of the entities or processes involved. (In fact, such a belief is, in the measurement case, impossible on pain of inconsistency and irrationality—as is shown by the insolubility proofs.) Thus the cognitive attitude that is needed to apply and use scientific fictions is pragmatically indistinguishable from belief, but it is not belief.
5. REPRESENTATION AND FICTION-BASED INFERENCE: AGAINST SIMILARITY AND ISOMORPHISM ACCOUNTS

Many of the most highly theoretical and mathematical models employed by scientists either involve fictional assumptions in their representations of real systems, or otherwise they refer, or purport to refer, to fictitious entities or processes. The representational character of these models can be explained roughly along the lines of a substitutive model of pictorial representation, which explains why the main cognitive function of these models is surrogate reasoning regarding its objects. This requires a “real- ist” attitude toward the model only in the sense of surrogate reasoning and inference, and as a powerful heuristic for further research. It does not require the belief that the model is an accurate representation nor that the entities described in the model are in fact real, and in fact it often will be inconsistent with such a belief.

In this final section I argue that this representational function of fictitious models in science is best understood by reference to a broadly construed “inferential” conception of scientific representation, such as the one proposed in Suárez (2004b). On this conception we can say of a model A that it represents an entity, system, or process B in case (i) the representational force of A is B, and (ii) A allows competent and informed agents to draw specific inferences regarding B. A virtue of this account is that it presupposes no relation of reference or denotation between A and B. The notion of representational force is defined so that it is fulfilled by any attempt at reference or denotation, however unsuccessful, that accords to the social practices and norms conventionally adopted in the use of such representational force. Also the notion of “inferential capacity” is fulfilled by any model that has sufficient internal structure to permit inferences regarding its “target,” regardless of whether it denotes it, or indeed regardless of whether it is intended to denote it.24

One of the advantages that I have claimed for this account is its ability to provide the most natural account of fictional representation.25 On this account there is no fundamental distinction in kind between a representation of a real entity or process, and a representation of a fictional entity (other than the obvious one of the existence of the entity so represented). Both are representations in virtue of fulfilling exactly the same conditions, and the nature of the representation is the same in both cases. Recall that according to this model there can be representation without reference: The existence of the target B is not required for conditions (i) and (ii) to obtain. In such cases we shall want to ask what the scientific interest of the inferences carried out under (ii) might be—and in most cases there will be inferences to true conclusions regarding observable aspects of the phenomena. Hence ether models allow a good deal of accurate predictions regarding the optical properties of light, and the plum pudding model allows, as we already saw, plenty of new and interesting predictions regarding atomic
observed phenomena; stellar structure models allow us to predict the correlations between the observable surface properties of the star, and so on.

In both the ether-based theories and the atom models cases, conditions (i) and (ii) are fulfilled. Maxwell’s theory was widely accepted to represent the electromagnetic configuration of an all-pervading medium capable of transmitting energy. As we saw, this was accepted by Maxwell himself, in spite of his rather skeptical views on the existence of the ether itself. The remarkable subsequent Cambridge scholarship in applying electromagnetic theory described in Warwick’s book demonstrates that condition (ii) is also satisfied. Huge amounts of effort, time, and energy are spent in the task to determine the inferential consequences of the ether models. Similarly, the plum pudding model is a model of atomic structure—and this is its representational target. Its empirical refutation by Rutherford shows conclusively that condition (ii) is also satisfied—because it shows that it is possible to test experimentally some of its consequences.

Stellar structure models satisfy (i) and (ii) trivially, because they represent real stars and allow very efficiently for inference regarding some of their properties. The quantum theoretical model of measurement interactions satisfies both conditions too—in spite of the internal inconsistency revealed by the insolubility proof. First, nobody doubts that it is a model, or representation, of measurement interactions at the quantum level—and second, condition (ii) is also fulfilled by the diverse applications (which I mentioned in section 2) of different combinations of assumptions from the quantum formalism in the different interpretations of quantum mechanics.

In addition, the inferential conception explains the substitutional character of fictive representations in science that I described in the previous section. The inferential conception denies that substitution is the essential constituent relation of representation, because it claims that no relation between A and B is required for representation. Hence it denies that “A is a substitute for B” is necessary and sufficient for A to represent B. However, the inferential account accepts that substitution à la Gombrich can be a means of the representational relation.26 That is, it accepts that this might be the mechanism or relation employed by scientists in surrogate reasoning about B on the basis of A.27 It thus explains why substitution plays a representational role in these cases.

The inferential conception was proposed in response to the deficiencies of other proposals for understanding scientific representation, namely, the isomorphism and similarity accounts.28 I believe we have here yet one more reason to adopt it in the face of its competitors, because neither isomorphism nor similarity can provide a similar explanation of the substitutional character of fictive representation in science. Take first the case of what we may call fictional representation, that is, representation of a nonexistent entity, such as the case of the ether. According to the isomorphism account, A can only represent B if they share their structure. But in the case of fictional representation this is either false or an empty truism, for it seems
impossible to ascribe structure to a nonexistent entity. If the ether does not exist it can not possess any real structure, so isomorphism can not obtain. There might arguably be also a sense in which the ether (a fictional entity) can be ascribed “fictional” structure, but in this sense it can not fail to have the structure that the theory ascribes to it, because it is fully defined by the theory. This points out a significant difference between fictional representations in science and elsewhere: The (fictional) properties of fictional entities represented by scientific models are implicitly defined by the models themselves, which need not always be the case with fictional representation in art or everyday life (for instance, some pictures of Santa Claus may be said to misrepresent him).

The same argument seems to apply straightforwardly to the similarity case. Suppose that everywhere there is about representation in science can be understood through the similarity account; that is, suppose that it were true that A represents B if and only if A and B are similar. Then if similarity requires the sharing of actual properties between A and B, there can be no “fictional” representation in science because B lacks any real properties. We cannot represent the ether by means of models that share actual properties with the ether if there is no ether. If on the other hand it is permissible to ascribe fictional properties to fictional objects and speak of a similarity between a real and a fictional object in virtue of a (fictional) sharing of actual properties—or an actual sharing of fictional ones—then similarity is automatic in cases of the fictions employed by science. It is built into the definition of the entity represented by the theoretical model that it (fictionally) has the properties ascribed to it in this model.

Now let us take the rather different case of the plum pudding model of the atom. What we have here is arguably a very different case of misrepresentation by a deficient model of a real entity simpliciter. Let us refer to this type of misrepresentation as fictive representation. The similarity conception fares a bit better here because it can be used to characterize the degree and respects in which the model differs from the entity modeled. (Version of the isomorphism conception can arguably do this too.) However, these conceptions can crucially not explain the heuristic power of the model, which seems by and large unrelated to any actual similarity or isomorphism between representational sources and targets. Of the three respects in which the model acts a heuristic guide to research (see section 2) these conceptions can only account for the third: Rutherford’s experimental refutation of the model. The other two respects in which the model proved heuristically powerful turned out to have nothing to do with any similarities between plum puddings and atoms, nor any isomorphisms between the mathematical structure of plum puddings and that of atoms.29

Mutatis mutandis, on either account, for the case of quantum measurement interactions. Whether there are genuine physical interactions
between microscopic entities and measurement devices might turn out to be a question of interpretation of the theory (in some versions of the many worlds or many minds theories the notion of genuine physical interaction arguably fails to play a role). In any case, similarity and isomorphism will not shed much light on the heuristic power of a fictional representation of a nonexistent process, nor on the fictive representation (i.e., misrepresentation) of a real one. These practices of fictive and fictional representation may at best be accommodated within these conceptions, but the heuristic value associated with such practices cannot be explained.

Yet in the inferential account, representation is neither false nor trivial in any of these cases. This is straightforward in cases of fictive representation. But then take the harder cases of fictional representation, such as Maxwell’s mechanical ether. In the inferential account, representational force could have failed to obtain (and indeed in the present-day, post-Einstein understanding of classical electromagnetism, this theory is no longer intended to represent features of the ether), and the models could have been mathematically so complex as to be completely useless for inference-making and surrogate reasoning—this would be the case for instance if the equations were not analytically solvable. Hence on the inferential conception the substitutive use of the virtuous fictional entities employed in science is a genuine achievement, which requires huge creativity, deep knowledge of the science, and a considerable cognitive ability in the use and development of sophisticated mathematical models.

6. CONCLUSIONS

The substitutive character of the fictions employed in science is best understood by means of an inferential conception of scientific representation. Although cases of fictional and fictive representation may be accommodated within the similarity or isomorphism accounts, the heuristic value of such kinds of representation cannot be convincingly explained. In other words, the isomorphism and similarity conceptions do not elucidate the distinction so dear to Vaihinger and to his present-day followers between scientific fictions and other kinds of fictions. According to Vaihinger this distinction can only be drawn at the level of cognitive function, and I have argued that the inferential conception of representation draws it precisely at the right level.

APPENDIX 1: THE INTERACTION FORMALISM

In this appendix I describe the tensor-product space formalism provided by the quantum theory of measurement to represent the interaction between
an object system and a measuring device. Given two Hilbert spaces, $H_1$ and $H_2$, we can always form the tensor-product Hilbert space $H_{1+2} = H_1 \otimes H_2$, with dim $(H_1 \otimes H_2) = \dim (H_1) \times \dim (H_2)$. If $\{v_i\}$ is a basis for $H_1$ and $\{w_j\}$ is a basis for $H_2$, then $\{v_i \otimes w_j\}$ is a basis for $H_{1+2}$. Similarly, if $A$ is an observable defined on $H_1$ with eigenvectors $\{v_i\}$ and eigenvalues $a_i$, and $B$ an observable on $H_2$ with eigenvectors $\{w_j\}$ and eigenvalues $b_j$ then $A \otimes B$ is an observable on $H_{1+2}$ with eigenvectors $v_i \otimes w_j$, and corresponding eigenvalues $a_i b_j$.

Consider two systems $S_1$ and $S_2$. If $S_1$’s state is $W_1$ on $H_1$, and $S_2$’s state is $W_2$ on $H_2$, we can represent the state of the combined system $S_{1+2}$ as the statistical operator $W_{1+2} = W_1 \otimes W_2$ acting on the tensor-product Hilbert space $H_{1+2}$. If either $W_1$, $W_2$ is a mixture, then $W_{1+2}$ is also a mixture. If, on the other hand, both $W_1$, $W_2$ are pure states then $W_{1+2}$ is pure. Suppose that $W_1 = P_{\psi}$ and $W_2 = P_{\phi}$, where $\psi = \sum_i c_i v_i$ and $\phi = \sum_j d_j w_j$. Then $W_{1+2} = \sum_i \sum_j c_i d_j v_i \otimes w_j$, which is a superposition of eigenstates of $A \otimes B$ in $H_{1+2}$. More specifically, if $S_1$, $S_2$ are in eigenstates of $A,B$, the combined system $S_{1+2}$ is in an eigenstate of $A \otimes B$. If $W_1 = v_i$ and $W_2 = w_j$, then $W_{1+2} = v_i \otimes w_j$, a so-called product state.

For an arbitrary (pure or mixed) state $W_{1+2}$ of the combined system, and arbitrary observable $A \otimes B$, the Generalized Born Rule applies. The probability that $A \otimes B$ takes a particular $a_i b_j$ value is given by:

$$\text{Prob}_{W_{1+2}} (A \otimes B = a_i b_j) = \text{Tr} \left( W_{1+2} P_{\psi} \right)$$

and the expectation value of the “total” $A \otimes B$ observable in state $W_{1+2}$ is:

$$\text{Exp}_{W_{1+2}} (A \otimes B) = \text{Tr} \left( (A \otimes B) W_{1+2} \right)$$

We will sometimes be given the state $W_{1+2}$ of a composite system, and then asked to figure out what the reduced states $W_1$, $W_2$ of the separated subsystems must be. Given a couple of observables $A$ and $B$ on $H_1$, $H_2$, there are some relatively straightforward identifications that help to work out the reduced states, namely:

$$\text{Tr} \left( (A \otimes I) W_{1+2} \right) = \text{Tr} \left( A W_2 \right)$$

$$\text{Tr} \left( (I \otimes B) W_{1+2} \right) = \text{Tr} \left( B W_2 \right)$$

where $I$ is the identity observable. This amounts to the demand that the probability distribution over the eigenspaces of observable $A$ (B) defined by the reduced state $W_1$ ($W_2$) be the same as that laid out over $A \otimes I$ ($I \otimes B$) by the composite state $W_{1+2}$, thus effectively ensuring that the choice of description (either in the larger or smaller Hilbert space) of a subsystem in a larger composite system has no measurable consequences as regards the monadic properties of the individual subsystems.
APPENDIX 2: THE FORMAL CONDITIONS OF THE QUANTUM MEASUREMENT THEORY

To formally describe the problem of measurement, we need to first introduce some notation, and denote by Prob(W, Q) the probability distribution defined by Prob_w(Q = q_n), for all eigenvalues q_n of Q. And let us denote Q-indistinguishable states W, W’ as W ≡ W’. Two states W, W’ are Q-indistinguishable if and only if Prob(W, Q) = Prob(W’, Q).

We may now enunciate the following two conditions on measurement interactions:

The Transfer of Probability Condition (TPC)

\[ \text{Prob} \left( U(W_o \otimes W_a)U^{-1}, I \otimes A \right) = \text{Prob} \left( W_o, O \right) \]

This condition expresses the requirement that the probability distribution over the possible outcomes of the relevant observable O of the object system should be reproduced as the probability distribution over possible outcomes of the pointer position observable in the final state of the composite (object + apparatus) system. (TPC) entails the following minimal condition on measurements employed by Fine (1970) and Brown (1985): A unitary interaction on a (object + apparatus) composite is a W_a measurement only if, provided that the initial apparatus state is W_a, any two initial states of the object system that are O-distinguishable are taken into corresponding final states of the composite that are (I \otimes A)- distinguishable. So we can use the pointer position of the measuring apparatus to tell apart two initial states of the object system that differ with respect to the relevant property.

The Occurrence of Outcomes Condition (OOC)

\[ U(W_o \otimes W_a)U^{-1} = \sum c_n W_n \text{ where } \forall W_n \exists \mu_n : \text{Prob}_{W_o}(I \otimes A = \mu_n) = 1 \]

This condition is often taken to express the requirement, inspired by the eigenstate–eigenvalue link, that the final state of the composite be a mixture over eigenstates of the pointer position observable. But to be precise, it expresses the more general idea that the final state of the composite must be a mixture over states in each of which the pointer position observable takes one particular value or other with probability 1.

NOTES

1. I am grateful for comments and reactions from audiences at the Fictions conference at Madrid’s Complutense University (2006), the APA Pacific
Division conference in Oregon (2006), the ZiF workshop at Bielefeld (2007), the Dubrovnik IUC conference (2007), and the workshop on scientific models at the University of Barcelona (2007). Thanks in particular to my commentators Carl Hoefer, Alfred Nordmann, and Eric Winsberg, and to Ron Giere for his detailed comments on a draft version of the chapter.

2. Vaihinger (1911, 1924).
5. For some examples see the essays collected in the volumes edited by Morrison and Morgan (1999), Magnani et al. (1999), Jones and Cartwright (2004), and de Chadarevian and Hopwood (2004).
7. This agrees with much of the recent literature on fictions in metaphysics and aesthetics—which tends to distinguish carefully fiction from falsehood. For instance, in Walton’s (1990) account of fiction as make believe, what characterizes fictions is their role as props in the prescription of imaginings, regardless of whether these turn out to be true or false. (Walton’s theory strikes me as inappropriate for scientific fictions for other reasons, which are not relevant to my discussion here.)
8. Fine draws a similar distinction between virtuous and vicious fictions (Fine, 1993, p. 5).
9. In the account that follows I am indebted to Morrison (2000, esp. chap. 3) and Warwick (2003, esp. Chapters 6 and 7).
10. See also Morrison’s extended discussion in this volume.
11. For instance, Trouton’s and Larmor’s ontological commitments to the ether could not have been more different. Trouton took second-order effects of the ether’s motion with respect to the earth to be measurable by ordinary optical instruments on the earth’s surface, whereas Larmor took any observable effects to be impossible as a matter of principle. See Warwick (1995).
13. My exposition is indebted to the essays in Buchwald and Warwick (2004), and in particularly to Chapter 1, Smith (2004).
15. Rutherford (1911).
17. Stellar astrophysics models provide a nice illustration of yet another thesis of mine, namely, that in scientific representation often the target is constructed along with the representation itself. I leave the detailed defense of this claim for another occasion (but see Suárez, 1999, for a preliminary account).
18. One of the earliest such proofs is due to Fine (1970). For details see Suárez (2004a).
19. Some of these options are described in Busch et al. (1991).
21. I agree with Lopes (1996, p. 79) that this requirement is too strong on substitutive accounts, and would yield contradiction.
22. This “substitutive” character of models has been noted in the recent modeling literature, particularly in the “mediating models” literature (Morgan & Morrison, 1999), where models replace their targets as the main “focus of scientific research.”
23. In fictive representation, the modeled entity (the target) is a real entity. In fictional representation, however, the target is also an imagined entity, just like the model.
24. Not all inferences will do, only those specific inferences that do not follow merely from the representational relation itself between A and B.
26. For more on the notions of “constituents” and “means” see Suárez (2003, 2004b).
27. I am assuming here that it makes sense to speak of a relation between an existing model and a fictional entity or process; the assumption that some intensional relations might lack the corresponding extensions is of course controversial, but it does not alter anything substantial in what follows.
28. The names of Ron Giere and Bas Van Fraassen are often associated to such accounts, but I believe the association is not altogether fair. Both Van Fraassen (1994) and Giere (2004) make it clear that intentional elements are also needed for representation. But I am not here arguing against straw men. Others have embraced such accounts on their behalf, and many others have assumed that the accounts were properly theirs.
29. To speak of the “mathematical structure” of concrete objects, such as plum puddings, is problematic in any case, because there is a well-known issue of underdetermination of any abstract structure by a concrete object or sets of objects.
References

260 References


References

References

Harris, T. (2003). Data models and the acquisition and manipulation of data. Philosophy of Science, 70, 1508–1517.
References


References


References

References 267


References


270 References


