THREE DOGMAS ON SCIENTIFIC THEORIES

Massimiliano Badino

Universitat Autònoma de Barcelona – Massachusetts Institute of Technology

DRAFT

ABSTRACT

Most philosophical accounts on scientific theories are affected by three dogmas or ingrained attitudes. These dogmas have led philosophers to choose between analyzing the internal structure of theories or their historical evolution. In this paper, I turn these three dogmas upside down. I argue (i) that mathematical practices are not epistemically neutral, (ii) that the morphology of theories can be very complex, and (iii) that one should view theoretical knowledge as the combination of internal factors and their intrinsic historicity.

---

1. INTRODUCTION AND STRUCTURE OF THE PAPER

Philosophy of science has been often concerned with, at times even obsessed by, scientific theories. As the crowning achievements of scientific research, theories have continuously lured epistemologists and methodologists. Or at least so it used to be. In recent years, the philosophical analysis of scientific theories has lost much of its appeal and it is now considerably less popular than it was in the past.[1] In this paper, I argue that the fading of interest in theories is related to the failure of the traditional accounts, which, in turn, is the result of three ingrained assumptions, or dogmas, on the nature and the role of scientific theories. I claim (i) that these assumptions have shaped virtually all philosophical views of theories, and (ii) that they prevent the elaboration of an account able to do justice to scientific practices present and past.

[1] For a discussion of this point see (Morrison 2007).

My argument consists of three steps. First, I discuss a general attitude in philosophy of science, which I call the Whewell thesis. This thesis expresses the cultural climate that allowed the three dogmas to thrive. Next, I proceed to elaborate more precise statements of the dogmas and to show how they permeate our philosophical views of theories. Finally, I bring forward my alternative. I argue that the philosophical project of proposing a full-fledged "theory of theories" is hopeless from the start. My suggestion is to frame a flexible workspace to use reflectively the theoretical knowledge already accumulated. By systematically overturning the dogmas, I hold that one should focus upon the epistemic cooperation between three theoretical dimensions: the representational, the transformational, and the explanatory dimension. My proposed workspace facilitates the conceptualization of the epistemic cooperation, but it does not provide for the details. A precise account of the production of theoretical knowledge is not a task for philosophical arguments alone, but it requires case-specific analyses combining cognitive science, psychology, philosophy of mind, epistemology, sociology, history of science, and so on.

Before starting my discussion, I need to lay out two provisos. First, here I am mainly concerned with theories in physics and chemistry. More generally, I focus upon theories that represent symbolically the world and deploy practices to manipulate these symbolic representations. Since an important part of my argument hinges on the role of mathematical practices, I deal with theories in which mathematics plays a substantial role. Second, my approach hinges on the fact that theories are knowledge producers i.e. they yield a particular kind of knowledge called theoretical knowledge. I suspect that practical knowledge—say the form of knowledge associated with the use of material tools, hand-on practices, or direct experience—can be included in my approach, but this is not my main goal at this stage.

2. The Whewell Thesis

One of my claims is that the three dogmas are the three-fold expression of philosophers' intellectualistic attitude toward theories. Traditionally, theories have been considered as objects of thought and a neat separation has been drawn between theories and their manipulations, applications, and specifications. I call this attitude the Whewell thesis. I choose to refer to William Whewell because he offers one of the tersest expressions of this intellectualism. His formulation will guide us to a precise statement of the dogmas.

The Whewell thesis lurks in some passages of his treatise on *Astronomy and General Physics* written for the Bridgewater series. The point of the whole book—and of the whole series of treatises, in fact—is to argue against the rampant French positivism and especially against the attempt at explaining nature without any divinity or divine designer. A step of Whewell's complex argument is to show that faith can cooperate with reason to produce genuine progress. From this perspective, the most revealing dualism is between Isaac Newton, the confident believer and Pierre Simon Laplace, the outspoken atheist. The superiority of the former over the latter follows from the different epistemological weight of their contributions to celestial mechanics. Newton is the discoverer of the laws of dynamics and the law of gravitation. Despite the fact that these laws "occupy little room in their statement", they contain all the knowledge that natural philosophy can possibly yield. Laplace, and other countrymen of his, extended the theory and applied these laws to the motion of the planets and their satellites. Although such extension is undeniably complex, requires a great deal of ingenuity, and at times it leads to epoch-making feats such as the "proof" of the stability of the solar system, still, it does not add a single grain of original knowledge:

> In explaining, as the great [French] mathematicians just mentioned have done, the phenomena of the solar system by means of the law of universal gravitation, the conclusions at which they arrived were really *included in the truth of the law*, whatever skill and sagacity it might require to develop and

> extricate them from the general principle. (Whewell 1837, 328 emphasis added)

While Newton gave us a genuinely new body of knowledge that illuminates the innermost secrets of the universe, Laplace and his colleagues merely unfolded the treasures incapsulated into the laws of dynamics and gravitation. Later in his book, Whewell stresses this point even more emphatically:

> When, therefore, we consider the mathematicians who are employed in successfully applying the mechanical philosophy, as men well deserving of honor from those who take an interest in the progress of science, we do rightly; but it is still to be recollected, that in doing this they are not carrying us to any higher point of view in the knowledge of nature than we had attained before: they are only unfolding the consequences, which were already virtually in our possession. (Whewell 1837, 332-333)

These passages of Whewell's book summarize an epistemological attitude that, I argue, has been more or less tacitly accepted by any philosopher who has reflected on the nature and the function of scientific theories. This attitude can be articulated in three theses or dogmas. The first dogma is the following:

> (D1) There is an epistemic asymmetries between the parts a theory consists in. More specifically, the whole knowledge content of a theory is stored in a handful of fundamental laws, while the mathematical techniques used to put these laws to work on concrete cases are epistemically neutral.

This point, which Whewell takes for granted, is the foundation of his whole argument. Once we have the fundamental laws, truth is "virtually in our possession" and the mathematical machinery just unearths what is implicit in the laws. There is no epistemic contribution of the mathematical practices. The second dogma reads:

> (D2) Theories present a core-belt morphology. In other words, we distinguish (i) a core of unrejectable statements that make the theory what it is and (ii) a set of auxiliary assumptions, concepts, methods, models, and techniques which can be sacrificed to increase the fitting between theory and experience.

The second dogma draws a clear line, as Whewell does, between those parts of the theory which are established once and for all and those parts that, although important, can be adjusted or even sacrificed for a better agreement with experiments. Thus, one of the reasons of Newton's superiority is that the laws of dynamics and gravitation will remain with us forever. On the contrary, in unfolding the knowledge packed in this theoretical core, Laplace and his colleagues introduced assumptions, concepts, techniques, and models that can be jettisoned without any traumatic blow for the theory itself if further experience forces us to do so. The third dogma is deceptively simpler than the previous two and can be formulated in the following manner:

> (D3) Theories are taken as a given.

Obviously, (D3) needs some elaboration. There are two major aspects of a scientific theory that a philosopher should account for: its internal structure and its temporal development. Ideally, a good philosophical conception of theories should inform us about how the various components of a theory stay together, how they depend on one another, and how they cooperate to produce knowledge. Additionally, since theories are human products subject to historical evolution, we also need to know how theories are modified, enlarged, accepted, disseminated, confuted, or rejected by different communities at different times. Both the structural and the historical aspects concur to illuminate that human artifact we call scientific theory. My claim is that traditional philosophical accounts drop systematically one or the other aspect and therefore treat theories, partially, as a given. Thus, dogma (D3) above can be explicated in terms of the two following theses:

> (D3.1) Philosophical accounts analyzing the internal structure of scientific theories, tend to overlook their temporal evolution and to consider them as a-temporal objects. They take theories as a historical given.
>
> (D3.2) Philosophical accounts analyzing the temporal evolution of scientific theories tend to assume a minimal internal structure—more precisely, the core-belt morphology of dogma (D2). They take theories as a structural given.

I submit that a satisfactory philosophical account of scientific theories should represent the temporal evolution of a theory as a structural object, the modification of the relations between its parts, the integration of new ingredients, and so on.

There are relations between the three dogmas. Dogma (D1) separates out what is epistemically fundamental and what is epistemically neutral in a theory. Thus, it concerns the distribution of knowledge within a theory. Dogma (D2) follows the same separation strategy, but introduces an intermediate grey zone. For instance, the laws of dynamics and gravitation represent the epistemic substance of Newtonian mechanics, while the mathematical techniques used to manipulate them do not add a single bit of knowledge. However, there are numerous assumptions, hypotheses, models, which enrich Newton's law in specific cases, but are not as essential. In the calculation of the motion of a planet it might be necessary to assume that its aphelion is immobile. This assumption does not inform us on the essence of the motion, but on its accidental circumstances. If the hypothesis does not help solve the problem, it can be dropped. In brief, (D2) deals with the internal structure of a theory. Thus, (D1) and (D2) combine to depict theories as structures composed of essential features, accidental features, and neutral mathematical practices.

The third dogma prescribes the way to approach the structure outlined by (D1) and (D2) and imposes a further distinction. It discriminates between the synchronic and diachronic analysis of the relation between the theoretical parts. The former is the investigation of the core and the belt in terms of their constituents and their relations. This investigation lays down a structural description that holds for any theory at any time.[2] By contrast, the latter is concerned with the way in which the belt changes over time given the unrejectable core. As a result, a diachronic analysis yields a description of the theory evolution over time that holds for any core and any belt.

[2] Note that it does not need to be a description of all theories ever produced. This analysis focuses upon what a theory ought to be i.e. the theory at its final stage of perfection.

3. Philosophical Accounts on Theories

3.1 Preliminary Considerations

Traditional philosophical accounts on scientific theories have interiorized the three dogmas (D1), (D2), and (D3). I do not have the space here for a systematic discussion of the entire philosophical literature. My argument consists in a swift survey and in displaying some textual evidence of the pervading presence of the three dogmas. Before starting, however, let me set the stage for this and the subsequent discussion.

I preliminarily distinguish between three kinds of epistemic jobs that we should expect performed by a theory. In the first place, a theory is supposed to give us a picture of the world populated with some sort of entities. One of the reasons why we frame theories is to represent our experience in symbolic terms. Mechanics paints the world as constituted by masses interacting by forces and possessing angular momentum; electromagnetic theory relies on charges, electric and magnetic fields, dipole moments; general relativity introduces even more exoteric entities such as gravitational waves and black holes. Any theory has its say on the furniture of the universe. Secondly, a theory is also supposed to solve concrete problems and to yield predictions. This is accomplished by manipulating symbolic representations mentioned above to produce new representations. Equations of motion must be solved, critical behaviors must be approximated, probabilities must be calculated and all these problems call for special mathematical techniques. Finally, a theory is requested to produce explanations of phenomena, i.e. to make sense of our experience. Statistical mechanics tells us that thermal equilibrium is a probabilistic event, general relativity interprets astronomical trajectories as space-time bending, electromagnetic theory presents optical phenomena as transverse vibrations of electric and magnetic fields. These are the three *dimensions* of a scientific theory, which I call representational, transformational, and explanatory.

These functions are carried out by the various constituents of a theory. For instance, fundamental laws such as the laws of dynamics, models, like the atomic model, ideal objects like the rigid body, concepts, like entropy, all concur to form a sort of picture about the external world. They tell us something about the furniture of the universe. Analogously, mathematical techniques, methods, procedures, tricks, algorithms allow us to manipulate our symbolic representations to produce applications to particular cases, predictions, and—generally speaking—other symbolic representations. Note, however, that a clear distinction between the epistemic dimensions of a theory does not mean a compartmentation of its constituents. While some theoretical constituents might serve *eminently* a function, they can also serve others at different times or in different contexts. For instance, the principle of conservation of energy tells us something about the world, but it might also be used as a calculation device. More importantly, I claim that theories produce knowledge by means of the epistemic cooperation between their constituents, i.e. there is a sense—to be explored later—according to which each constituent contributes to each dimension.

An additional point to bear in mind is that my account aims at conceptualizing how historically concrete theories are assembled, work, and eventually change. Hence, I am not concerned with theories as abstract logical objects, but as a historical and social event embedded in a web of texts and practices. From this perspective, Clairaut's theory of the Moon is given by all the ingredients appearing in the papers in which Clairaut develops his theory of the Moon. This is the reason why, in my account, Clairaut's theory of the moon is not only the laws of dynamics and gravitation, but also all the mathematical techniques he uses.

3.2 Permanence of the Dogmas

For the time being, I postpone the discussion of the explanatory dimension and concentrate on the other two. Traditionally, the issue of explanation has been separated from the problem of the nature of a scientific theory and, temporarily, I will follow this usage. In this section I want to argue that the dogmas introduced above permeate philosophical reflections on the nature of scientific theories. I do not want to make a specific point in history of philosophy of science, though. This would require much more space and a much deeper discussion. I view the dogmas (D1)-(D3) much like religious dogmas. As such, they might have been held with shifting tenacity by different philosophers or even not held all at the same time. But they characterize by and large the philosophical approach to theories and their presence can be seen lingering over most of the influential views. At the beginning of the next section I offer another argument for the fundamentality of these dogmas. Here, I confine myself to a cursory overview of the main accounts.

The acceptance of (D1) among philosophers of science is fairly uncontroversial, with some qualifications which I will discuss in a moment. Typically, philosophers do not see mathematical practices as making any epistemic difference and the reason is not hard to see. In fact, most philosophers do not view those practices as a theoretical constituent at all. The—often implicit—argument goes that mathematical methods are imported into physical theory from pure mathematics and thereby adopted as a mere external aid to inference. A natural consequence of this argument is that, since mathematical practices are external to physical theories, they reveal deep connections between diverse fields and serve the purpose of unification (Morrison 2000). Thus, mathematical structures help establish relations between theories, but do not directly contribute to the epistemic content of the theory they feature in. Incidentally, besides being historically questionable, this argument does not entail the epistemic neutrality of mathematical practices.

Furthermore, the acceptance of (D1), does not mean that philosophers have unanimously held that fundamental laws are the unique storage of the knowledge content of a theory. To be sure, the challenge of this common wisdom has carved out a moderate epistemological space for some mathematics. Ronald Giere, for instance, laments that philosophers have often missed the important role played by approximations in scientific practice:

> Those scientists and philosophers who have taken it for granted that the laws of nature are well confirmed, general statements have obviously not been ignorant of the fact that scientists regularly use approximations. But they have taken this to be a relatively inconsequential fact about science. They have regarded the fact as a matter of only practical, not theoretical importance. (Giere 1988, 78)

Against this habit, Giere claims that fundamental laws are only schemes from which approximate models of physical systems are constructed. Approximations, therefore, are part and parcel with the process of producing knowledge. Nancy Cartwright pushes this line of argument even farther. She claims that fundamental laws are explanatory (in the usual D-N sense) to the extent that they are general and therefore literally not true of any special case. To give a truthful analysis of the phenomena, one needs to supplement fundamental laws with assumptions, approximations, and hypotheses that make the resulting phenomenological laws highly problem-oriented (Cartwright 1983).

These reflections show a clear intention to go beyond the epistemic dominance of the fundamental laws, but still fall short of challenging the first dogma. In the first place, approximation techniques contain much more than just mathematics. In fact, it seems that Cartwright refers more specifically to those models and assumptions that represent the particular system under study. To apply Boltzmann's transport equation to a certain fluid is not a matter of just mathematics, but also requires assumptions on the behavior of the fluid. Furthermore, the mathematics of a theory is not limited to approximation techniques, but includes all the practiced used to manipulate symbols. Coordinate transformations or Fourier integrals are examples of these practices although they are not immediately useful to approximate fundamental laws.

The core-belt morphology described in dogma (D2) is closely related to what Dudley Shapere dubbed the *inviolability thesis*: "the idea that there is something about the scientific (or, more generally, the knowledge-seeking or knowledge-acquiring) enterprise that cannot be rejected or altered in the light of any other belief at which we might arrive, but that, on the contrary, must be accepted before we can arrive, or perhaps even seek, such other beliefs" (Shapere 1984, xix-xx). According to Shapere, the inviolability thesis drives most philosophical investigations on theories. Logical Empiricism assumes this thesis when claims that there is a handful of propositions playing the role of axioms from which the other propositions of the theory can be logically derived.[3] Logical techniques, like mathematical practices, are epistemically neutral. In the Semantic View, this idea is still central. The upholders of this approach[4] regard a theory as a collection of models but they still consider some models as basic exemplars. As Giere shows, mechanics textbooks are organized around fundamental models such as the pendulum or the linear oscillator. These are the exemplars that must successively be enriched to represent real physical systems.

Furthermore, the core-belt morphology is a pillar of virtually all history-oriented accounts of scientific theories. The reason is clear enough: according to this morphology, the evolution of a theory over time is a combination of tradition—the untouchable core of the theory—and transformation—the refinement of the auxiliary assumptions. In this way, it is possible to talk about the *same* theory changing in time or, alternatively, a sequence of *different* theories belonging to the *same* research program. This is the rationale of Lakatos' approach (Lakatos 1978), and also Larry Laudan has followed basically the same strategy, although he allows for a certain rearrangement in the constituents of the core:

> During the evolution of any active research tradition, scientists learn more about the conceptual dependence and autonomy of its various elements; when it can be shown that certain elements, previously regarded as essential to the whole enterprise, can be jettisoned without compromising the problem-solving success of the tradition itself, these elements cease to be a part of the "unrejectable core". (Laudan 1977, 100)

---

[3] See for instance (Suppe 1977; Giere 2001; Mormann 2007).

[4] Here I refer especially to (Suppe 1989).

Finally, dogma (D3). In the early decades of the 20th century, it was customary to treat theories as articulated bodies of linguistic and/or nonlinguistic objects kept together by relations of some formal nature. The fact that theories are historically situated entities in a socially organized world was supposed to play no role in philosophical analysis. The actual evolution of theories of the past was largely ignored, while theories of the present, such as relativity and quantum mechanics, had to be explicated in terms of axioms, theoretical/observational language, correspondence rules, models, isomorphism and so on. Thus, Newtonian mechanics was regarded as made of the three laws of dynamics, the law of gravitation and their logical consequences or, alternatively, of the models in which those laws are true. It was of no epistemological importance that those laws were not understood by Newton the way in which a 20th century philosopher of science would. Regardless if one sees theories as axiomatic systems of propositions or holds that "to axiomatize is to define a set-theoretical predicates" (Suppes 1969), the philosophical analysis of the theoretical internal structure entails a pertinacious tendency to separate neatly between what a theory formally boils down to and how a theory came into being (Reichenbach 1938).

There are good reasons for this tendency. A first reason is that by isolating a theory as a well-defined object, philosophers are following a venerable scientific practice. Like scientists, they single out the relevant aspects of the phenomenon under study. Another reason is that the aim of philosophical inquiry is to elucidate what turns a set of propositions, models or other into a scientific theory irrespective of when and where the theory happens to be formulated. A 'secret sauce' for a scientific theory cannot be time-, place-, or user-dependent. Sometimes this argument is phrased in terms of the irreconcilability between philosophical normativeness and historical descriptiveness. In support of the normative project, Ronald Giere has added that, at best, philosophy should be concerned with contemporary science: history is philosophically irrelevant (Giere 1973).

The advent of a more historically-minded philosophy of science in the late 1960s only moves the "giveness" from without to within, from the context to the structure. Theories were endowed with a historical dimension at the expenses of their internal architecture. This is a consequence of the fact that, with the entering of history, the analysis moved up to a higher level. Philosophers lost interest in theories as such and began to inquiry larger entities such as paradigms, research programs or research traditions that spanned longer periods of time. Logical Empiricism and the Semantic View analyzed scientific theories in terms of their constituents and investigated the manner in which they are combined together. By contrast, since the 1960s, philosophers began to favor the view of theories as "answers to problems". This evanescent conception makes theory parasitic on the notion of problem. Instead of analyzing the theoretical structure, philosophers chose to produce a taxonomy of the possible (empirical and conceptual) problems and then to characterize theories as the corresponding answers.[5]

But there is a deeper reason. To unfold it, let us consider for a moment a recent example. Olivier Darrigol argued that physical theories possess a "modular structure", i.e. they are constituted of building blocks, or moduli, which are, in themselves, established pieces of theoretical work (Darrigol 2008). In creating a theory, scientists may use bits of Hamiltonian mechanics, classical optics, thermodynamics, statistical physics and so on, all arranged in a coherent structure. Darrigol's view helps understand how theories are accepted and disseminated, but does not illuminate much their internal structure from an epistemological standpoint. In fact, the moduli are themselves theoretical bits taken as givens. On the other hand, Darrigol's is one of the best attempt a philosophically-minded historian of physics could try. The reason is contingent upon the development of philosophy of science in the 20th century. Among the upholders of Logical Empiricism and its immediate emanations, it was commonplace to hold the view that theories are accepted, disseminated, and learned because of internal, structural reasons. To put it a bit bru-

[5] See for instance (Laudan 1977; Shapere 1984, 286-287).

tally, a theory is believed by its practitioners because is true, elegant, simple, well-behaving, or just working. By contrast, if people believe a false, clumsy, abstruse theory, this must be due to some external factor such as ideology, social or personal interests, economic advantages, political agendas, and alike. History and sociology can—and ought to— explain biases, deviations, and errors, while believing in truth does not need further explanation. The philosophy and history of science of the 1960s detected the fallacy behind this claim and moved resolutely to the opposite direction. They denounced internal factors as positivistic naivety and treated theories as a structural given. My point is that this radical reversal is as little beneficial to our overall understanding of theories as the strict internalism of Logical Empiricism.

Even this quick discussion of the main philosophical accounts should suffice to conclude, with Dudley Shapere, that "[t]here is today no completely—one is almost tempted to say remotely—satisfactory analysis of the notion of a scientific theory" (Shapere 1984, 112). My goal in the remainder of the paper is to provide an alternative. I argue that we have to think about theories in a way that allows us to combine their being social and historical activities with an analysis of their internal structure, which can produce some general, epistemological insights.

4. A Working Space of Theories

4.1 The Dispensability of the Dogmas

The previous survey shows that philosophical accounts seek for a difficult equilibrium between rigorous arguments—often reached by formal methods—and the necessity to capture actual scientific practices in the present and the past. Initially, philosophers favored approaches based on logic or set theory able to lead to formally correct conclusions. Those conclusions, however, resembled very little what scientists actually do. This criticism paved the way to historical- and cultural-oriented approaches in which the emphasis moved on tacit knowledge, hand-on practices, paradigms and so fort. The three dogmas are the result of the early phase in which formal methods ruled undisputed. However, when in the 1960s philosophers began to relax some of the formal constraints and to pay more attention to history of science, they did not abandon the three dogmas. In fact, the history-oriented philosophy of science reconfigured them to ultimately debunk the internal theoretical structure. The price paid to maintain the dogmas was a philosophical withdraw from the ambition of describing the architecture of theories. By contrast, my contention is simply that we should restore this ambition and discard the three dogmas.

There is another possibility. One could instead discard the ambition of mirroring actual scientific practices and hold a strong normative stance for philosophy. In this case, the argument would go that one should distinguish neatly between describing scientific practices and providing norms concerning what science ought to be. These are two different cultural projects. The latter is philosophy's main goal and problems crop up when it is mixed with the former, which instead belongs to history of science. However, one moment's reflection shows that this argument relies on a false dualism. One could draw a clearcut line between normative and descriptive projects only if it were possible to found the former either (i) on a set of a priori principles of rationality or (ii) on transcendental arguments. For, if one is not allow to derive scientific norms from actual scientific practices, then either they are imposed as a priori rules of rationality or it is proved that science cannot happen without them. Both strategies seem unfeasible at the moment. A more positive argument against the strong normativeness is the following. One should notice that scientific practices in the past and in the present enjoy a quite respectable amount of success. This seems to suggest that, at least partially, they approach the ideal of what a theory ought to be. From this perspective, the normative project is not independent of—let alone incompatible with—the description of historically located scientific practices, but it can build on them to elucidate their functioning and to ameliorate their efficacy. This form of normativism is also compatible with my project of replacing the three dogmas with more suitable theses about theories.

4.2 Beyond the Dogmas

The popularity of the three dogmas has deeper roots than one can think at first sight. A closer investigation reveals that they are expressions of very natural philosophical attitudes. The first dogma, for instance, embodies the usual strategy of characterizing the object of study in terms of its essence, what makes it what it actually is. The theoretical core is the theory's innermost identity. Analogously, the picture of theory evolution as the refinement of auxiliary components respect to an unrejectable core embodies the model of the classical ontological accounts of becoming. Finally, the third dogma defines the subject matter of study as a well-bounded object—even if only for mere analysis' sake an attitude that, as I argue above, is rooted in scientific tradition.

The problem with these natural attitudes is that theories are very peculiar objects. Theories are devices to produce a certain form of knowledge—i.e. theoretical knowledge—and philosophical accounts ultimately aim at obtaining theoretical knowledge about them. But the meta-theories hitherto proposed do not differ substantially from theories themselves. The study of theories is therefore prone to the dangers of reflexivity. What is more, philosophers have tackled this issue is a way that amplifies these dangers. They determine what a theory should look like to infer how it should work. This simple consideration contains the ultimate kernel of the inadequacy of the three dogmas. They lead to a hopelessly circular approach because one should have already a conception of what a theory is to elaborate a theoretical account of theories. In setting the epistemological arrow from essence to functioning, philosophical accounts objectify scientific theories and strip them of the only characterizing trait of which we have direct understanding: their being producers of a specific kind of knowledge.

The cumbersome circularity implicit in the three dogmas suggests that we should look at the issue from another perspective. We do not know what a theory is, but we do have a considerable amount of *theoretical knowledge*. We should therefore invert the epistemological arrow and try to clarify what a theory is through its being a knowledge producer instead of the other way round. To do that, we cannot simply construct a theory of theories, but instead we have to look at theories from the perspective of the practices that constitute them. This change of perspective would allow us to use reflectively the theoretical knowledge at our disposal to illuminate the functioning of theories. The procedure has two steps. First, we have to set up a *workspace* to conceptualize how the internal structure of a theory produces knowledge. Contrary to the usual philosophical accounts, this workspace does not aim primarily at telling us what a theory ultimately is. Instead, it serves as a platform of epistemic interaction between the various constituents of a theory. As I show below, it provides a minimal and very provisional answer to the question "what is a theory", but this is not its main purpose. In fact, the workspace leaves open many questions about the specific way in which the theoretical constituents interact and the consequences of this interaction for the production, acceptance, and dissemination of knowledge. The second step, therefore, consists in using the theoretical knowledge at hand in various fields to enrich the picture. Briefly said, the workspace makes possible a reflective interaction of our theoretical knowledge to illuminate the process of production of theoretical knowledge. Let us see how to flesh out this proposal.

To begin with, recall my previous claim that a theory consists of three dimensions: representational, transformational, and explanatory. These dimensions represent the parts of the epistemic job of knowledge production: symbolically representing the external world, transforming these representations to solve specific problems, and producing explanations of the observed behaviors. I now submit three new theses to replace the dogmas. In the next section I will discuss more thoroughly their consequences. My first thesis is the following:

(1) No theoretical constituent is epistemically neutral in principle.

This claim means that each dimension contributes to the overall knowledge content of the theory or, at least, is not in principle excluded from such contribution. Note, however, an important point. Thesis (1) does not mean that each dimension works separately from the other and that their contributions add to the knowledge content. This view is compatible with the epistemic neutrality of the mathematical practices, which is contained in the first dogma. What (1) really means is that elements of different dimensions serve eminently different purposes, but *they may also partially contribute to the other dimensions*. Thus, for instance, certain mathematical practices does not merely serve the purpose of transforming symbolic representations, but they can also—at times—provide a representational content themselves. For this reason, I call (1) the *thesis of the epistemic cooperation*. Note also that (1) does not entail that each dimension ought to contribute to the others all the time. How the cooperation between dimensions occur depends on specific cases and, ultimately, on the choices of the theoretician. The precise dynamics of the epistemic contributions is historically situated.

The second thesis deals with a similar question and reads as follows:

(2) Theoretical constituents do not have a preconceived arrangement.

This claim is naturally related to (1). While (1) removes the notion of epistemic neutrality, (2) removes the center-periphery distinction, i.e. the morphology of the theory is more variegate than the usual core-belt structure. To detail this morphology—which can change from theory to theory and from time to time—one has to shift the focus from where the knowledge is stored to how it is produced. It is the specific manner in which the theoretical dimensions cooperate that establishes what elements are more important, how they are arranged, and how they contribute to the overall knowledge content. I call (2) the *thesis of the complex morphology*.

Finally, the third thesis:

(3) The process of production of theoretical knowledge depends (i) on the structure of the theory and (ii) on how this structure evolves in time and place.

The idea behind this thesis is that not only the production of knowledge depends on a situated epistemic cooperation between the dimensions and on a situated morphology, but also the constituents of the dimensions are historically, socially, and culturally situated. Thus, (3) tells us that we cannot reduce the knowledge content of a theory either to its internal structure, or to its specific history and cultural context. In fact, these two aspects interact because we cannot understand the epistemic cooperation between the dimensions without taking into account the historical determination of their constituents and we cannot understand their historical evolution without taking into account how they cooperate. I call (3) the *thesis of the two-sidedness of theoretical knowledge*.

Before discussing the workspace in more details, let us pause a moment to note the difference between theses (1)-(3) and the traditional accounts. Following the dogmas, philosophers usually made a series of reductivist choices. Either they choose to investigate the nature of a handful of fundamental laws, or they characterize theory evolution as the combination of something that remains and something that changes. Theses (1)-(3) points at a completely different direction. The historicity of a theory is built into the theoretical constituents themselves. Theories are cultural and historical entities not because they are immersed in a certain "context" and because some of their parts are replaced over time, but because they produce knowledge through the epistemic cooperation of historically and culturally situated practices. Furthermore, theses (1) and (2) suggest that also the way in which they cooperate and their organization is itself historically and cultural situated. Taken together, theses (1)-(3) claim the intermingle of structural and historical factors in knowledge production.

4.2 Cooperating Dimensions

Theses (1)-(3) define the general features of the workspace, but they still leave undetermined the details of the epistemic cooperation between dimensions. My contention is that these details cannot be established by means of abstract philosophical arguments, but must be obtained from the theoretical knowledge at our disposal. However, we can further elaborate the concepts of theoretical dimensions and epistemic cooperation.

The representational dimension of a theory concerns its ability to depict the physical world by selecting relevant features and by expressing them symbolically. In other words, it creates a *symbolic codification* of the phenomenon under study. A codification is typically performed by means of concepts, principles, models, ideal objects, and alike. Thus, the motion of earth, sun, and moon is codified by the three-body model, the concepts of mass, velocity, force, orbital elements, the principles of conservation of linear and angular momentum and so forth. It should be remembered, however, that these theoretical entities do not need to be exclusively representational. The transformational dimension concerns the manipulation of symbolic codifications to obtain other symbolic codifications by means of *symbolic practices*. The immediate purpose of these transformations is the solution of specific problems. Consider, to continue our example, the actual calculation of the relative motion of earth, sun, and moon. To predict the positions of the three heavenly bodies anytime in the future, one needs to formulate the equations of motion, to simplify them, possibly by coordinates transformations and integrals of motion, and to solve them by approximations or numerical methods. All these steps require specific mathematical techniques to turn a general symbolic codification into the calculation of the concrete behavior of a planet or a satellite.

Finally, a theory must also explain phenomena, a function that goes beyond both representing the world and solving concrete problems. I have postponed the discussion of the explanatory dimension because it has a somewhat different status. My claim is that the explanatory quality is a feature of the theory *as a whole*.

There is no space in this paper even for a sketchy overview of the literature on explanation. At any rate, my goal here is not to provide a full-fledged theory, but only to lay out some very general requirements. To begin with, I take the explanatory dimension of a theory to give causal explanations.[6] I do not exclude that also other accounts, such as the unification account, can be included in this framework.[7] I submit that a theory yields an explanation of a phenomenon when it gives an *epistemic story* of this phenomenon with two crucial characteristics. First of all, it must be a *causal* story, a requirement that can be decomposed in two further conditions.

> (EX1) The theory gives a set of causal factors in terms of symbolic codifications of the features of the phenomenon.

Conditions (EX1) states that the explanatory theory exhibits a representation of the features of the phenomenon that act causally, i.e., in Woodward's terms, they are the variables that we need to manipulate if we want to provoke a causal difference in the phenomenon. The second condition reads:

> (EX2) The theory gives a mechanism relating the causal factors with the phenomenon.

To explain, a theory needs to tell us how to modify the causal factors to obtain a certain change in the ensuing phenomenon. Conditions (EX1) and (EX2) jointly give us a causal structure (in Woodward's sense) and tell us *why* and *how* phenomena happen. However not all the why-and-how explanations are also *good* explanations. How good an explanation is does not entirely depend on satisfying conditions (EX1) and (EX2). As pointed out by thesis (3), the knowledge content combines structural and historical factors. I therefore add the following social-epistemological requirement:

---

[6] On the theory of causal explanation I follow especially (Woodward 2003). According to Woodward's manipulability account, the variable *X* codifies a causal factor for the quantity codified by the variable *Y* if there is a possible intervention on *X* such that we can change the values of *Y*: "putting this in the form of a slogan, we can say that manipulability accounts are committed to the following: *No causal difference without a difference in manipulability relations, and no difference in manipulability relations without a causal difference*" (Woodward 2003, 61).

[7] In fact, I do not think that the causal theory is incompatible with the unification approach; see, for instance (Strevens 2004).

(EX3) A good explanation for a community is an epistemic story that satisfies a set of historically, socially, and culturally situated requisites of the community.

The acceptability of an explanation is a community-dependent property. Hence, it depends on conditions that the community has interiorized as part of its history and culture. To clarify this point I introduce another important concept. I have claimed that an explanation is an epistemic story of the phenomenon to be explained. As any narrative, an epistemic story is a story of *a certain kind*. In other words, it is organized around one of more *epistemic tropes* that establish what cognitive resources to mobilize, what concepts, mathematical techniques, and argumentative patterns to deploy to satisfy (EX1) and (EX2). Epistemic tropes are the constituents of the explanatory dimension. They might be methodological directives, scientific principles elevated to the rank of metaphysical norms, generalized concepts. Epistemic tropes characterize different kinds of explanatory stories. Mechanistic, evolutionary, field-theoretical, historical, functional explanations differ by the resources used, by the argumentative patterns adopted, by their social and cultural impact, all elements that depend on the epistemic tropes. It is now clear that (EX3) is a condition about the epistemic tropes. The rationale to accept, develop, and disseminate an explanation changes from community to community and from time to time in relation to the conceptual and technical resources applied and the argumentative patterns used to align them. My point is therefore that part of the reasons to accept, interiorize, and disseminate a theory comes from the kind of epistemic story it tells: the standards to select cognitive resources, the argumentative patterns to align them, the inclusion or exclusion of possibilities, the reading of the meaning of the problem.

The next step in my argument is to show how symbolic codifications, symbolic practices, and epistemic tropes cooperate to produce theoretical knowledge. Let us begin with the interrelation between representational and transformational dimensions. As we have seen in the foregoing sections, philosophy of science has traditionally neglected the epistemic contribution of what I call symbolic practices.

Traditionally, they have been deemed as epistemically neutral. By contrast, my thesis (1) claims that they possess a substantial knowledge content beyond their ability to transform a symbolic codification into another. I suggest that symbolic practices contain an *epistemic surplus* that yields a certain amount of knowledge about the problem on which they are applied. The intuition behind this suggestion is that, as tools, symbolic practices must fit both our cognitive capabilities and the symbolic codifications they act upon. This double *fitting* tells us something very general about the object of the transformation as well as about our cognitive capabilities, a kind of knowledge intrinsic in the use of the tools. For instance, the successful application of convergent trigonometric series to a mechanical problem might suggest that the problem has a periodic nature. This information is part of the fact that those mathematical practices help solve the problem. It is, roughly speaking, another aspect of that fact.

Note that, because of the *double* fitting the epistemic surplus incapsulated in the symbolic practices is not univocal and a priori, but it is always socially and historically situated. Different actors can view the same symbolic practice in different ways at different times or emphasize one of the end of the fitting at the expenses of the other. Probabilistic tools are a notorious case in point. For some scientists, the successful use of probabilistic methods to solve a problem indicates that the phenomenon is intrinsically stochastic, whereas for others it can indicate a limitation of our knowledge of the phenomenon. At times, the difference in reading the epistemic surplus depends on historical cases. An example is the different receptions of special relativity. Einstein saw his mathematical techniques as telling something really important on simultaneity and the individual judgements of the order of physical events. Instead, for Herman Minkowski all that mathematical machinery revealed rather the deep geometrical blending of space and time and for Ebenezer Cunningham it represented a clever extension of the treatment of electrostatic problems to a time-dependent situation.[8]

A second and partially related point is that symbolic practices reveal a body of knowledge that is not included in the symbolic codification. They yields a very general information on the nature of the problem in terms of the conditions for its manipulability. This difference in the body of knowledge conveyed is the basis for the epistemic cooperation between representational and transformational dimensions. To treat a theoretical problem, one has to negotiate between the type of information obtained by using certain concepts and certain practices until a satisfactory combination is reached. The result is a *center of structured practices and codifications* that deals with a specific problem in the overall architecture of the theory.

A physical theory can contain several such centers, each serving a different function. They constitute the series of steps through which the theory solves partial problems to obtain its final goal. The way in which they are arranged is the morphology of the theory and, as stated in thesis (2), it can assume complex forms. This bring us to the issue of the cooperation with the explanatory dimension. The epistemic trope shapes the explanatory story of the theory by laying down argumentative patterns into which symbolic codifications and practices can be aligned. Thus, the explanatory dimension makes sure that the various centers are coordinated to produce a consistent story of a certain kind. But it is important to realize that the construction of an epistemic story is itself the result of a bidirectional negotiation. To see why, consider that epistemic tropes can be expressed by different argumentative patterns. For instance, if we want to assemble an explanation of a thermodynamic phenomenon based on the idea of strict irreversibility, we can either prove that the reverse phenomenon is physically impossible, or we can define an entropy

[8] The point of the concept of epistemic surplus is to insist on the relation between a body of knowledge and the act of symbolic transformation. This does not mean that a certain knowledge content is invariably related with a certain practice. Quite the opposite, because as the examples show, the epistemic surplus must be situated into historical and cultural traditions where it can acquire different forms. Again, the specification of these forms is a task for historical, cultural, psychological, and epistemological analysis.

function and show that it increases monotonically. These are two different argumentative patterns—incidentally, both adopted by Planck at different stages of his radiation theory—that embody the same epistemic trope. As a consequence, the realization of an epistemic trope involves a continuous negotiation between the argumentative patterns, the representational, and the transformational resources.

Furthermore, symbolic codifications and practices effectively influence the construction of the epistemic story as stated by the condition (EX3) above. Concepts, models, assumptions, and techniques have an important role in making an explanation good or bad for a certain community. They actively determine a crucial component of the theoretical explanation. As an example, take the clash between upholders and critics of the Copenhagen interpretation of quantum mechanics. The heart of the disagreement mainly concerns the use of certain concepts (e.g. hidden variables) and/or mathematical techniques and the interpretation of their consequences on the overall epistemic story of quantum phenomena.

This discussion of the epistemic cooperation leaves many questions unanswered. How should we conceptualize the epistemic surplus? How and when is it incapsulated into practices? What are the mechanism of negotiation between symbolic codifications and practices? What kind of morphologies are possible? On what conditions does the epistemic dimension drive the choice of concepts and practices and on what conditions it is constrained by them? The larger point of this philosophical project is that these and similar questions must be addressed by using the theoretical knowledge in our possession. Only the interaction between disciplines such as epistemology, cognitive science, psychology, sociology, social epistemology, philosophy of mind, history of science and cognate fields can help us understand how theories are created, accepted, communicated, and disseminated. Theories present themselves as complicated alignments of symbolic codifications and practices. Very much as in the case of cultural systems, one cannot hope to understand scientific theories from without. It is necessary to apply knowledge from different sources and to be immersed in specific cases to unfold the epistemic cooperation between dimensions, the morphology, the several intersections and interrelation, and their change over time. A theory must be deconstructed and deciphered using other theoretical knowledge. My workspace should serve as a post-dogmatic platform to ease this deconstruction.

5. Conclusion

The main point of this paper is that traditional accounts of theories, based on the three dogmas, have forced philosophers to choose between the analysis of the internal structure and the description of the historical evolution. My claim is that this choice can be avoided by changing our way to look at scientific theories.

In the second part of the paper, I have supported the view that theories are to be taken essentially as knowledge producers. We should therefore concentrate on the process of generating knowledge as the result of an epistemic cooperation between three dimensions of the theory: representational, transformational, and explanatory. I argue that we cannot establish what a theory is or ought to be by means of purely philosophical arguments. We need to use the wealth of theoretical knowledge at our disposal to clarify how theoretical knowledge actually works.

The picture of theories emerging from this account is very different from the traditional one. Here theories are not well-defined and strongly structured assemblages of propositions or model. Instead, theories live in a web of texts and practices, their boundaries are porous, their internal morphology is ever-changing. They are cultural products whose epistemic dynamics, and therefore their normativity, can be illuminated only by keeping in mind their inherent historicity.

References

Cartwright, Nancy. 1983. How the Laws of Physics Lie. Oxford: Oxford University Press.

Darrigol, Olivier. 2008. "The Modular Structure of Physical Theories." Synthese no. 162:195-223.

Giere, Ronald. 1973. "History and Philosophy of Science: Intimate Relation or Marriage of Convenience?" British Journal for the Philosophy of Science no. 24:282-297.

–––. 1988. Explaining Science: A Cognitive Approach. Chicago: University of Chicago Press.

–––. 2001. "Theories." In Blackwell Companion to the Philosophy of Science, edited by William H. Newton-Smith, 515-524. Malden, MA: Blackwell.

Lakatos, Imre. 1978. The Methodology of Scientific Research Programmes. Edited by John Worral and Gregory Currie. Vol. I, Philosophical Papers. Cambridge: Cambridge University Press.

Laudan, Larry. 1977. Progress and its Problems. Berkeley: University of California Press.

Mormann, Thomas. 2007. "The Structure of Scientific Theories in Logical Empiricism." In The Cambridge Companion to Logical Empiricism, edited by Alan Richardson and Thomas Uebel, 136-162. Cambridge: Cambridge University Press.

Morrison, Margaret. 2000. Unifying Scientific Theories: Physical Concepts and Mathematical Structures. Cambridge: Cambridge University Press.

–––. 2007. "Where Have All the Theories Gone?" Philosophy of Science no. 74:195-228.

Reichenbach, Hans. 1938. Experience and Prediction. Chicago: University of Chicago Press.

Shapere, Dudley. 1984. Reason and the Search for Knowledge. Dordrecht: Reidel.

Strevens, Michael. 2004. "The Causal and Unification Approaches to Explanation Unified—Causally." Nous no. 38 (1):154-176.

Suppe, Frederick. 1977. "The Search for Philosophical Understanding of Scientific Theories." In The Structure of Scientific Theories, edited by Frederick Suppe, 3-232. Chicago: University of Illinois Press.

–––. 1989. The Semantic Conception of Theories and Scientific Realism. Chicago: University of Illinois Press.

Suppes, Patrick. 1969. "Models of Data." In Studies in the Methodology and Foundations of Science. Selected Papers from 1951 to 1969, 24-35. New York: Springer.

Whewell, William. 1837. Astronomy and General Physics Considered with Reference to Natural Theology. London: William Pickering.

Woodward, James. 2003. Making Things Happen: A Theory of Causal Explanation. New York: Oxford University Press.