Scientific Coordination beyond the *A Priori*

A three-dimensional account of constitutive elements in scientific practice

PhD dissertation of Michele Luchetti

Department of Philosophy

Central European University

Supervisor: Prof. Maria Kronfeldner

Associate Supervisor: Prof. Flavia Padovani

Date of Submission: December 18th, 2019

Table of Contents

Abstract

In this dissertation, I present a novel account of the components that have a peculiar epistemic role in our scientific inquiries, since they contribute to establishing a form of coordination. The issue of coordination is a classic epistemic problem concerning how we justify our use of abstract conceptual tools to represent concrete phenomena. For instance, how could we get to represent universal gravitation as a mathematical formula or temperature by means of a numerical scale? This problem is particularly pressing when justification for using these abstract tools comes, in part or entirely, from knowledge which is not independent from them, thus leading to threats of circularity. Achieving coordination between some abstract conceptual tools and the concrete phenomena that they are supposed to represent is usually a complex process, which involves several epistemic components. Some of these components eventually provide stable conditions for applying those abstract representations to concrete phenomena. It is in this sense of providing certain conditions of applicability that different philosophical traditions, as well as some contemporary reappraisals, view these components as *constitutive* or *a priori*.

In this work, I present a new gradualist, contextualist, and relational approach to understand these constitutive components of scientific inquiry. It is gradualist inasmuch as the degree to which some component is constitutive depends on three quantifiable features: quasi-axiomaticity, generative potential, and empirical shielding. Since the quantification of these three features impinges on the history and practice of using these components in a scientific context, my approach is a contextualist one. Finally, my approach is relational in a double sense: first, it identifies ordinal relationships among epistemic components with respect to their constitutive character; second, these relationships are relative to a scientific framework of inquiry.

After introducing my account and a classic example of constitutively *a priori* principles, i.e., Friedman's (2001) analysis of Newtonian mechanics, I turn to my own case studies to demonstrate the advantages of my approach. Firstly, I discuss Okasha's (2018) view of endogenization as a pervasive theoretical strategy in evolutionary biology and suggest that the constitutive character of the core Darwinian principles progressively increases with endogenization. Secondly, I apply a conceptual distinction between two varieties or scopes of coordination – general coordination and coordination in measurement – to Ohm's work on electrical conductivity. This distinction allows me to pinpoint to what extent components along different dimensions (e.g., instrumentation, measurement, theorising, etc.) were constitutive of the forms of coordination which Ohm relied on. Thirdly, I discuss the epistemic function of the Hardy-Weinberg principle in the history and practice of population genetics. I assess this principle in terms of my account and identify approximation and stability as two components that are highly constitutive, in that they contribute to justifying its use in population genetics.

Finally, applying my account to these case studies enables me to identify at least three qualitatively different types of constitutive components: domain-specific theoretical principles, material components, and domain-independent assumptions underlying reasoning abilities. In the light of my results, I draw some general conclusions on epistemic justification and scientific knowledge.

I hereby declare that:

- This dissertation contains no materials accepted for any other degrees in any other institutions.
- This dissertation contains no materials previously written and/or published by another person, except where appropriate acknowledgment is made in the form of bibliographical reference, etc.

December 18th, 2019

Michele Luchetti

Introduction

> The roads by which men arrive at their insights into celestial matters seem to me almost as worthy of wonder as those matters in themselves.
>
> (Johannes Kepler, *Astronomia Nova*)

I. The knower, the known, and the Kantian revolution

Science is contested for many reasons. One of them is that it provides us only with partial, sometimes incompatible, and often disorienting pictures of our world. One source of this criticism lies in how scientists deal with the gap between the world and the descriptions of it that they produce, a gap that is intrinsic to the task of representing a concrete world by means of abstract concepts. Scientists try to fill this gap by *coordinating* the abstract tools they use with the concrete phenomena they aim at describing. In doing this, they rely on various presuppositions, whose justification may sometimes seem unclear, weak, or absent. Despite all this, many of us would not hesitate to say that several scientific inquiries resulted in some of the most reliable and useful knowledge that our species ever gained about the world and ourselves.

From the very beginning, the active search for reliable knowledge of the world went hand in hand with the reflection on the status and character of this knowledge. The way in which scientists and, more generally, societies conceive of the connection between the knowing subject and the world is at the root of most questions concerning the character of scientific knowledge, the role of the knower, and the effort itself of trying to understand their relationship. In fact, some view of the relationship between the subject and the object of (scientific) knowledge seems to be in the background of most of our intellectual efforts, including scientific inquiries and philosophical reflections. Sudden as well as stepwise changes throughout the history of science and philosophy radically transformed our standpoint on this relationship several times. This is a central reason for investigating the ways by which scientists get to reliable knowledge of the world and the features that make some knowledge scientific.

What has been called Kant's 'Copernican revolution' was one of the crucial turning points in the history of western thought, and it had a lasting impact on our conception of the relationship between the knower and the known. One of the questions Kant addressed was: "how is natural science possible"? In his view, answering this question required delving deep into the relationship between the knowing subject and the knowable object. According to Kant, our knowledge of the world emerges from a functional relationship between our knowing faculties and our experiences. By 'functional' he meant that any knowledge is the outcome of an interaction between two actors, the knowing subject and a knowable object, where this outcome is irreducible to its components, since the subject necessarily transforms the object in order to represent it and know it. Our empirical knowledge should not be conceived as the result of simple abstractions from sense data, which are available to us without mediation, as empiricist thinkers would (roughly) put it. Any object of empirical knowledge is an object of knowledge *for us*, in the sense that our knowing faculties structure in certain ways our experience of the world. In other words, we can know about the world only what *we can* know about it. This does not mean that, in Kant's view, we construct the world or that the world is mind dependent. It simply means that we know the world only inasmuch as it can be grasped by our knowledge systems which, given our limited human faculties, are themselves also limited.

Kant's move had important epistemological implications. It was deemed as a revolution in the sense that it overturned the view that the world is something immediately given to us, which the knowing subject can boundlessly approach in order to find its very 'nature'. In Kant's opinion, metaphysics should not deal with whatever transcends the knowable empirical world, since we could not possibly access that world, but it should operate within the limits of human reason and its knowing faculties. It is only within these limits that the world can present itself to us as a possible object of knowledge. As a result of the revolution, also the meaning of 'objective' changes. According to Kant, our knowledge of the world is not objective in the sense that it corresponds to or approximates a state of affairs as it is *per se*, independently of our knowing faculties. In no way could we test whether this sort of correspondence holds, because this would require an independent, absolute way to access the world epistemically. This sort of access is precluded to us limited beings. This Kantian stance is echoed by those contemporary philosophers of science holding that any representational system's capacity to represent cannot hold in virtue of any reference to a reality (or portion of it) completely external to the system itself. Therefore, the standards of objectivity by which we can evaluate any scientific description of objects are dependent on the character and structure of our representational devices. At this very general level, objectivity, for Kant, must be understood as the conformity of the object of knowledge to the necessary, universal, and fixed prescriptions of the knowing subject – the human knowing faculties, not the individual subject – so that that object of knowledge presents itself to us in the way it does.

This revolutionary conception of the relationship between knowing subject and knowable world had several effects – at times visible, at times hidden – on the subsequent history of scientific and philosophical thinking. At various points of this history, most pillars of Kant's theory of experience and knowledge were harshly criticised or dismissed in the light of empirical discoveries. Yet, this did not suppress the influence of the intuitive core of his view, almost a sentiment: the sentiment that the knowledge of the world that we produce with our scientific systems unavoidably bears our human signature. This dissertation is about some recent reappraisals of one of the harshly contested pillars of Kant's theory of knowledge and experience: his notion of *a priori*.

According to Kant, certain structures of the human mind play a role in our cognitive processes, in such a way that they, along with some of the resulting knowledge claims, can be regarded as *a priori*. However, to understand his notion of *a priori*, we cannot directly connect it to the meaning it had in the rationalist and empiricist philosophical traditions, which identified the root of knowledge, respectively, in the intellect and in the senses. In the rationalist tradition, the intellect was considered the only source of certain knowledge because it can achieve true judgments entirely *a priori*, that is, before the messy realm of our sensations enters the stage. According to rationalist thinkers, the certainty of deductive methods, warranted by the neat workings of reason, was the antidote against the scepticism that arises from the inevitable mistakes of our faulty sensory organs. In opposition to this view, the empiricist tradition admitted only one legitimate path to true knowledge, that is, our sensations. Faulty as they may be, our sensory organs are the only means by which we can gain knowledge about the world, while the workings of our intellect – a *tabula rasa*, when we are born – are enabled by the abstraction of concepts out of bundles of sensations. By the term *a priori*, both traditions referred to knowledge claims the truth of which could be achieved without recourse to sensory input.

In Kant's epistemological system no priority is accorded to any of these two sources of knowledge. On the one hand, our faculty of sensibility is the source of perceptions, which are possible in virtue of two general *a priori* conditions of experience, space and time. Space and time are, according to Kant, functions of the mind, innate mental faculties that enable us to structure our perceptions and make sense of them as of something out there in the world. In other words, for something in the world to be a possible object of experience for us, it must be structured according to spatial and temporal coordinates. On the other hand, our faculty of intellect is the source of the logics of reasoning by means of which we organise the sensible material into knowledge judgments. These logics of reasoning, or *a priori* categories, are the formal conditions of our thinking and, according to Kant, they mirror the judgments of formal (Aristotelian) logic. For something to be a possible object of our cognition and perception, it must be expressed according to the formal rules prescribed by these categories. This is the sense in which space, time, and the categories are *a priori* for Kant, as structural preconditions for our experience and knowledge of the world.

Here, it is not necessary to delve into Kant's own distinction between different kinds of *a priori* knowledge judgments, the analytic and the synthetic. In the first chapter of this dissertation, I will get back to this distinction and to why it is important for contemporary reappraisals of Kant's notion of the *a priori*. Yet, I must emphasise that the identification between the *a priori* understood as constraints intrinsic to our cognitive faculties and the *a priori* as a set of propositions – identification promoted by Kant's own ambiguous examples – had a substantial impact on subsequent traditions of interpreting the *a priori*.

II. Updating the Kantian *a priori*: naturalised, cultural, and epistemic traditions

Let us step back for a moment. Kant's epistemological turn contributed to making sense in a systematic way of the incredible leap forward in the understanding of the world that humankind had recently experienced with the historical progression that went from Copernicus, through Kepler and Galileo, to Newton. His view of the joint work of sensibility and intellect as our faculties of experience and cognition seemed to overcome the dead ends of rationalism and empiricism. At the same time, Kant's system also aimed at securing the future path of scientific knowledge by uncovering the limitations of our knowing faculties. This critical move was not meant to constrain scientific progress, which would still be able to unfold in innumerable ways, but to avoid it transcending the limits of human reason and ending up in some metaphysical aberration. Yet, Kant's most lasting insight is that knowledge – more specifically, scientific knowledge – is produced by operating within certain rules of the game which, in his view, are the rules prescribed by the *a priori* conditions of our experience and cognition. To our contemporary minds it sounds rather striking that philosophy should have the task of unveiling and investigating those conditions. Should not science itself, if anything at all, have this duty? The subsequent history of philosophical and scientific thinking shows that the sciences progressively eroded philosophy's right to identify the limits within which science itself does, or should, operate. Yet, this is also a history of how different sciences sometimes competed and sometimes worked in parallel, with respect to one another and to philosophy itself, in the effort to pinpoint some of these conditions and limitations. Psychology, history, anthropology, and sociology tried to identify something that resembled the Kantian *a priori* while, at the same time, progressively thinning down philosophy's impact on the subject.

Kant's system aimed at uncovering the ultimate metaphysical boundaries of scientific knowledge. However, the view that space, time, and logical categories are fixed, universal, *a priori* conditions for experience and cognition soon started to be considered untenable. A few decades after Kant's publication of the *Critic of Pure Reason*, that same scientific knowledge advanced in directions that were thought to contradict the *a priori* character of those structures (more on this in chapter 1). Soon after, a scientific earthquake forced humankind to rethink its position in the universe and, as part of this, the knower-known relationship. Although evolution by natural selection had certainly its most pervasive and traumatic impact on how we had to reconceive the place of the human species in the natural world, it also impacted the relationship between knowing subject and knowable world. The position of our species lost its place of centrality and discontinuity with respect to rest of the natural world and found its new position in one among the many branches of the tree of life. This relativization also meant that our cognitive faculties were a product of evolution and, as such, nothing but a result of adaptive processes.

Natural selection endowed us with a brain with potentialities far exceeding the immediate demands that our species had hundreds of thousands of years ago. The human reasoning faculties are not special, since they arose from the very same process which endowed many other species with their perceptual and cognitive skills. Yet, the sheer fact that the evolutionary novelty of the human brain stored so much potential, in comparison to the adaptive needs of our new-born species, generated a huge gap that seems to constitute a *unicum* in the history of life. This paradox is so striking, especially when considering how long it took for our species to develop its capacity to put its cognitive faculties to use. Our basic brain configuration has been stable for at least a hundred thousand years, whereas the rate of our mental, or cultural, evolution spiked only in the last few thousand years, and its pace started increasing exponentially with the progress of modern science. Observed from this angle, the velocity with which scientific reasoning developed intuitively seems to be at the antipodes of a view of human knowledge and experience bounded by fixed categories, which, as I mentioned, collapsed under the influence of empirical discoveries brought about by scientific progress itself. In fact, while the progress of the sciences seemed to require what, in Kant's view, were fixed conditions of experience and cognition to change, the biological evolution of the brain could not explain these changes, at least with respect to our astonishing cultural and scientific progress. At what level are *a priori* conditions for experience and knowledge to be found, if any at all? It is on the gap left between the biological limits of human cognition and the tortuous but increasingly rapid paths of scientific progress that the many scientific and philosophical endeavours to update the view of the *a priori* focused their attention.

One fruitful direction was the attempt to naturalise the Kantian *a priori*. Scientists and philosophers with empiricist leanings looked for a way out from the strictures of Kant's epistemological system by interpreting his *a priori* in terms of biological preconditions of experience. Famously, Helmholtz developed a theory of human sense-perception in which physiological generalisations, rather than Kant's transcendental arguments, provided access to the necessary and universal conditions of human experience.[1] In the light of Helmholtz's theory, "[s]cience could now be seen as an open system of knowledge: a totality which is constantly growing and changing as a result of experience [...] instead of delivering truth via fixed categories and intuitions, science is understood as a gradual approximation of truth" (Bitbol et al. 2009: 9). In general, the rise of Darwinian thinking fuelled the naturalisation of the *a priori* along two different directions. In a rather indirect way, the application of evolutionary reasoning to culture and knowledge opened the way to its understanding in terms of a selective process which constrains the development, diffusion, and survival of theories and ideas. Although the understanding of this selective process as a kind of *a priori* was certainly not explicitly in the mind of most evolutionary thinkers, this mutual influence was almost certainly operating somewhere under the surface (cf. section 2.3.1 in chapter 2). In a more direct way, the naturalisation of the *a priori* proceeded along the physiological path, with the intention of reducing the *a priori* to an evolutionary product of natural selection in the biological sense. The long echo of this route reaches contemporary cognitive science and developmental psychology, with researchers emphasising the role of innate concepts in shaping and constraining our experience (Carey 2009), and neuroscientific research, where neurocognitive mechanisms like predictive coding are explicitly investigated in line with the Helmholtzian tradition (Friston & Kiebel 2009).

The psychological or physiological interpretation of the Kantian categories was rejected by those philosophers and scientists who maintained that the genuine meaning of the *a priori* had to be that of a set of preconditions for all empirical knowledge, including those parts of our biology and psychology that investigate our perceptual and cognitive organs. Therefore, those preconditions could not be identified by means of biological or psychological empirical generalisations. In other words, the *a priori* structures were not to be found in our organic limitations to the acquisition of empirical input, but in those cultural or epistemic conditions that justify and constrain the ways in which we structure that empirical input. As we saw, Kant's universal and fixed *a priori* had been shown to be inadequate in the light of the growth of empirical knowledge. As an upshot, the only viable way to update it was to make it dynamic, that is, to relativize it to those conditions that provide the ultimate justification for our

[1] Cf. especially Helmholtz (1867).

empirical generalisations, be these conditions epistemic, historical, social, anthropological, pragmatic, or any combination of the above. Accordingly, the focus of interest must shift from the development of the human species and the biological basis of its perceptual and cognitive faculties, to the development of our culture – and, in particular, of our science.

This direction was followed by different philosophers, who variously emphasised either the epistemic or the cultural character of the *a priori*.[2] The Marburg school of neo-Kantianism, especially with Cassirer, famously historicized the epistemic interpretation of the *a priori* in the light of the new scientific developments that Kant's static view could not accommodate. A parallel attempt was pursued by Reichenbach, who later abandoned these neo-Kantian leanings, as other fellow logical empiricists, such as Carnap, had done. In other scientific and philosophical milieus, the pragmatist and conventionalist traditions reconceived the epistemic *a priori* according to interpretations more in line with their concerns. Later in the 20th century, Kuhn's work and the subsequent wave of social studies of science stimulated the emergence of a more distinctively cultural interpretation of the *a priori*, which added a sociological dimension to the relativization of the epistemic *a priori*. In this sense, the *a priori* was identified with those social structural conditions that enable and shape the production of scientific knowledge, including the epistemic and non-epistemic values endorsed by scientists and epistemic communities.

III. Contribution of this dissertation and summary of chapters

In this dissertation, I only consider a small subset of the several reinterpretations of the *a priori*, the origins of which I summarised in this introduction. Contemporary reinterpretations are so dispersed across different disciplines, including developmental psychology, cognitive neurosciences, philosophy of science, social studies of science, cultural anthropology, etc., that it hardly makes sense to consider them in a relationship of family-resemblance anymore. My starting point, in this work, is the branch housing the contemporary reappraisals of the *epistemic* interpretation of the *a priori*.

The contribution of this dissertation is threefold. First, this work provides support to an epistemological constructivist stance in philosophy of science. To achieve this, I analyse various contemporary reinterpretations of the epistemic *a priori*, which rarely communicate with one another and lie scattered even in the restricted domain of philosophy of science, and I structure them as branches of a unitary debate. My purpose, in this sense, is to show the connections and increase the exchange among views that sometimes do not seem to recognise themselves as kin. Then, I raise

[2] For some context, see, for instance, Friedman (2000a).

concerns with respect to the coherence, usefulness, and aims of reappraising the epistemic interpretation of the *a priori* in general, to show the methodological strengths of this move. Finally, I situate this move in the context of the needs and current trends in philosophy of science, to offer a glimpse of the benefits of this epistemological stance for current science and philosophy.

The second contribution is a more constructive one. A large part of this dissertation is dedicated to the development of an original account of *constitutive elements*, i.e., my own reinterpretation of the epistemic *a priori* as preconditions for scientific inquiry. My account takes inspiration from both contemporary and older reinterpretations, and from other philosophical sources that are not directly linked with discussions of the *a priori* in scientific knowledge. Yet, the perspective I develop is a novel one, mainly with respect to two aspects. On the one hand, my account is based on understanding the *constitutivity* of certain epistemic components as a degree-sensitive property that depends on different contextual factors. What counts as constitutive is, therefore, not simply relative to an epistemic framework, as in most contemporary views, but it is also a matter of degree. On the other hand, I partly shift the emphasis of the debate away from scientific theories, i.e., the epistemic dimension towards which dynamical views of the epistemic *a priori* have directed most of their attention. My account is influenced by the focus on science in practice of many contemporary philosophers of science and, yet, it avoids downplaying the importance of theories. Inspired by what may be viewed as an ecumenical spirit, I eventually identify three distinct kinds of constitutive elements along different epistemic dimensions, including those of theorising, measurement, experimentation, and instrumentation.

Finally, I support my account by discussing three case studies, in which I combine insights from the history and practice of science. Even though the main purpose of these cases is that of demonstrating the advantages of my account, some insights may be regarded as valuable in and of themselves. More specifically, I investigate two cases from the biological sciences and one from the physical sciences. In the chapters dedicated to these cases, the reader will find some original contributions concerning the interaction between instrumentation, experimentation, and theorising, the epistemic role of certain principles in the history of biology, the analytic categories to deploy in epistemology of measurement, and the history of the electrical sciences. In what follows, I provide a breakdown of the chapters.

Chapter 1 will begin with an introduction to the transcendental perspective in philosophy of science as opposed to objectivist views of science and provides a brief history of the epistemic interpretation of the *a priori* in parallel with the development of philosophy of science as a discipline. After presenting Quine's challenge to the analytic/synthetic distinction and a few counterarguments, I will outline the main contemporary interpretations of the epistemic *a priori* and discuss how some features of the Kantian *a priori* (necessity, universality, fixity) are weakened or transformed in these perspectives. Finally, I will introduce and defend the background commitments of my own approach, including the integration of history and philosophy of science, the attention to scientific practice, and the pluralism with respect to the epistemic dimensions of scientific inquiry to be considered.

In chapter 2, I will develop a novel degree and context-sensitive approach to constitutive elements in science. The chapter will start from a discussion of the notions of 'constitutive', 'coordination', and 'presupposition', in order to outline a working definition of 'constitutive' that is in line with the methodological commitments presented in chapter 1 while, at the same time, preserving the core meaning of this notion as it figures in reappraisals of the epistemic *a priori*. After that, I will introduce in detail the most influential of these reappraisals, namely, Friedman's view of the 'relativised *a priori*'; I will present two of the core examples from the history of science supporting his view; and I will outline some criticisms that show the limitations of his approach. Finally, after presenting some useful conceptual tools from evolutionary epistemology, I will introduce the three main features of my account of constitutive elements in science: quasi-axiomaticity, generative potential, and empirical shielding. To show the advantages of my approach, I will reframe one of Friedman's own examples, and compare my account to other epistemological perspectives.

In chapter 3, I will present a case study from the history of evolutionary biology to demonstrate the usefulness of my account for scientific cases that usually are not considered in discussions of the constitutive *a priori*. More specifically, I will examine Okasha's view of endogenization as a strategy of theorising in evolutionary biology and I will argue that it increases the constitutivity of the core Darwinian principles. To do that, I will show how attempts at formalising the core structure of the theory of natural selection fostered endogenization, which is initially characterised by the functional extension of the core Darwinian principles, and later by the semantic extension of some core conceptual resources of the theory of natural selection. In other words, I will suggest that current discussions in evolutionary biology concerning the *meaning* of some concepts, such as 'inheritance', express the attempt at re-coordinating the core Darwinian principles with the novel empirical domains to which the epistemic *function* of these principles had already been extended. Finally, I will argue that the degree of constitutivity of these principles, expressed by the three features of my account, increased with the use of endogenization as a strategy of theorising.

Chapter 4 will develop the claim that there are kinds of constitutive elements that are different from the theoretical ones on which most of the literature on the epistemic *a priori* focuses, and which concerns the material conditions of scientific inquiry. To support this claim, I will discuss a case study that provides contributions in both the epistemology of measurement and the history of the electrical sciences. Firstly, I will suggest a conceptual distinction between 'general coordination' as the general issue of coordinating abstract representations with concrete phenomena discussed in chapters 1 and 2, and 'measurement coordination', which concerns the issue of how quantity terms acquire meaning through measurement. Then, I will provide a historical reconstruction of how the physicist Georg Ohm contributed to a novel form of measurement coordination between some electrical quantities and the procedures to measure them. I will suggest that this achievement was itself a necessary, although not sufficient, precondition for Ohm to formulate his famous law. After that, I will discuss the role of the material apparatus used by Ohm and its constitutive character with respect to the coordination he relied upon. Finally, I will generalise the discussion on materiality, to show systematically that scientific instruments belong to another epistemic dimension, complementary but alternative to the theoretical one, in which certain elements can have a constitutive character.

In chapter 5, I will suggest that a third kind of constitutive elements can be characterised as domain-independent, that is, as functioning beyond the limits of specific scientific frameworks. By means of a case study in population genetics, I will illustrate how some domain-independent preconditions underlie reasoning abilities that are deployed in several scientific inquiries independently from the specific empirical domain under investigation. The case will focus on the epistemic role of the Hardy-Weinberg principle in the history and practice of population genetics. By means of my account, I will show that this principle has a certain degree of constitutivity, in that it represents a counterfactual ideal state of equilibrium that enables the representation of concrete phenomena, in this case, real changes in genetic frequency distributions, as deviations from expected frequencies. Yet, I will emphasise that the use of this principle is itself justified by two further constitutive elements: approximation and stability. I will outline the epistemic function of the principles of approximation and stability and argue that they have a different constitutive character than domain-specific principles. Finally, I will suggest that they are domain-independent principles underlying reasoning abilities used across scientific frameworks.

In sum, this dissertation will show that the coordination between the abstract representational tools used by scientists and the concrete phenomena that they are supposed to represent is usually the result of a complex process. This process may comprise several 'layers' of coordination, that is, several levels between the abstract and the concrete, and it can involve many epistemic dimensions, including that of theorising, experimentation, measurement, and instrumentation. Some epistemic components operating along different layers and dimensions of this process have a peculiar justificatory character, in that they provide some conditions for the applicability of the abstract representation to the concrete phenomena. It is in this sense that these components should be considered as constitutive. As I will argue, their constitutivity is a matter of degree and context, and these components cannot be viewed as qualitatively homogeneous.

CHAPTER 1

1.1 Situating a transcendental approach within an empiricist philosophy of science

1.1.1 Transcendentalism and naturalism in philosophy of science

What are the sources of our scientific knowledge of the world? How is it possible for us to obtain this knowledge from experience? Do our ways of engaging the world epistemically, that is, our modes of knowing, represent mere *access points* to the empirical reality? Or is there anything about the human element in scientific knowledge that makes the latter non-reducible to its empirical components, and thus comes to constitute our scientific picture of the world?

Certainly, doing science means taking the empirical input seriously; so seriously that the relatively short history of our sciences has gone through many radical changes in the light of new empirical findings. However, offering a philosophical perspective on the scientific picture of the world means to take seriously also how these changes took place and what they tell us about the relationship between our modes of knowing and the natural world. In other words, it means to give justice to the complexity and variety of ways of reasoning and epistemic tools that we developed along the way. To do this, a safer and more epistemically humble starting point consists in considering our ways of knowing as *lenses* through which we investigate reality, rather than mere access points. Yet, these lenses prove to be quite hard to factor out, in the moment we want to take stock and consider the natural world as independent from our interactions with it. Investigating the structure, shape, and qualities of these lenses is integral to the effort of trying to understand how it is possible for the world to present itself to us in the ways it does.

I begin with a preliminary step. Here are three common philosophical standpoints that – I suggest – should be temporarily bracketed, in order to start any fruitful analysis of the epistemic conditions of scientific inquiry:

1. **Assumption of identity between acceptance and truth**. There are a lot of instances in which scientists accept and use theories or models even though they do not believe them to be, strictly speaking, true (whatever we mean by 'true'), but because their acceptance is advantageous for more context-specific and goal-oriented reasons (van Fraassen 1987, Wimsatt 2007). A philosophical investigation of the epistemic conditions of scientific inquiry should not take for granted that truth *in the abstract* is itself the only goal that drives the activities and development of science.

2. **Assumption of truth as direct correspondence to states of affairs.** This point connects with the previous one. If we want to account for how our scientific inquiries capture the natural world as successfully as they do, we cannot presuppose that a direct correspondence between our scientific representational tools and phenomena in the world holds independently of those very tools and our relationship with them. Considering this correspondence as something 'given' in the absolute sense, as entirely independent of our epistemic activities, leads to the claim that: "we can investigate and represent the world in the ways we do simply in virtue of the fact that the world is such and such". However, this rarely seems a satisfactory answer.[3] The possibility of conceiving truth as correspondence to states of affairs depends on a range of factors that allow us to investigate and represent a state of affair through certain means and for some purposes. Yet, this consideration does not mean that conceiving truth as correspondence is unjustified *tout court*. Even less does it want to embrace the idealist twist that states of affairs hold in virtue of our scientific representations being such and such. It only emphasises that truth as correspondence cannot be assumed as independent from certain conditions that justify that correspondence. Truth as correspondence is something that is achieved from within a framework of inquiry, and not as an absolute given.
3. **Assumption of naïve realism**. If taken seriously, points 1 and 2 should lead to embrace a temporary bracketing of the realist assumption of a world of ready-made phenomena, that we can capture independently of both our ways of engaging with it and our representational systems.[4] This bracketing is functional to explaining the process itself of investigating and representing the world scientifically.

The last point expresses a transcendental stance, by which I mean the suspension of the belief that the world discloses itself to us independently of the ways we engage with it, in order to investigate, instead, the *conditions of possibility* for us to investigate and represent the world scientifically. A transcendental move of this sort is usually contrasted not only with a naïve realism, but also with epistemological naturalism, i.e., the view that "the study of science is itself a scientific enterprise" (Giere 1988: 12). According to naturalists, there is nothing that lies in principle outside the scope of scientific explanation. Thus, a transcendental approach seems – to most naturalists' eyes – a bold attempt to attribute to philosophy the privileged function of explaining science, under the assumption

[3] Cf., for instance, Block (1983), Giere (2010), Goodman (1968), van Fraassen (2008).
[4] Cf. Massimi (2008).

that the latter is an object of knowledge which cannot be itself explained scientifically.[5] To counteract this possible objection, I would like to emphasise that a transcendental approach in philosophy of science should simply be conceived as the temporary bracketing of existence claims about what we normally consider as scientifically explainable, in order to focus on *how it is possible* that it can be explained scientifically. To answer the questions "How is it possible that phenomenon x can be accounted for scientifically?" and "What is required for phenomenon x to be explained scientifically?", certainly we can – and should – make use of empirical data (sociological, historical, cognitive, etc.). Yet, the mode of organisation and justification of this empirical input, the emphasis on framing the *explanans* as a *condition of possibility*, and the act of bracketing itself are certainly philosophical in nature. These are aspects of the investigation and reflection *on* science as an object of inquiry that elevate this reflection at a meta-level with respect to the science it accounts for.[6]

In the rest of this section, I will provide a cursory historical sketch of how one of the most influential philosophical traditions of the 20th century, the scientific philosophy of the logical empiricists, turned the Kantian *a priori*, which he viewed as conditions of possibility for empirical knowledge, into linguistic/propositional elements, and how this reverberated in the broader philosophical debate up to the end of the century. In section 1.2, I will outline the major reappraisal of the transcendental *a priori* in Kantian terms by contemporary analytic philosophy, i.e., Friedman's view of the relativised *a priori* principles, and the alternative stances that, within this approach, have been developed against the background of his view. Finally, in section 1.3, I will motivate the goal of this dissertation in the context of the current state of the art, and I will outline some methodological points that guide my approach.

1.1.2 A brief history of the constitutive *a priori* in the 20th century

A transcendental approach was typical of neo-Kantian perspectives on science in the late 19th and early 20th centuries influencing the early work of many logical empiricists, such as Carnap's doctoral thesis (*Der Raum*, 1918), and Reichenbach's *Habilitation* thesis (*Relativitätstheorie und Erkenntnis Apriori,*

[5] Cf. Callebaut (1993: 1-5, and specifically footnote 3) on the varieties of naturalism and their relationship with a transcendental stance.

[6] Here, I see my goal as consonant with Chang's appeal for humanism in epistemology, even though he focuses on a rather different point: "[…] instead of taking 'natural' as something transcending humans and seeking forever to erase all traces of humanity in our concepts, we can embrace the humanity of our concepts and assess their merits in terms of how well they enable human scientific inquiry" (Chang 2015: 43). Cf. also Chang (2020).

1920).[7] Central to the transition from the neo-Kantian tradition to logical empiricism was the issue of how to characterise the 'third' kind of knowledge beside *a posteriori*, empirical knowledge and *a priori* knowledge, devoid of empirical content; that special kind that Kant had labelled the *synthetic a priori*. According to Kant, certain principles, or functions of thought, can extend our knowledge without incorporating new empirical input from the world. For instance, the proposition that 'all events have causes' is not established inductively, nor is it an outcome of purely logical analysis. It is a precondition of experience that extends our knowledge without any novel empirical data being considered.

Kant argued that also space and time are not features of the world that we can know empirically, but rather general preconditions of our experience of the world and, in this sense, synthetic *a priori*. Following Cassirer (1923: 269), the correct understanding of the Kantian *a priori* is that by which "a cognition is called *a priori* not in any sense as if it were *prior* to experience, but because and in so far as it is contained as a necessary premise in every valid judgment concerning facts".[8] This means that, by understanding space and time as *a priori*, Kant conceived them as general functions of our mind (more precisely, of our intuition), to which our experiences conform. Yet, many later philosophers thought that his view of space and time as *a priori* meant that the propositions of Euclidean geometry and Newtonian mechanics were *a priori* knowledge and, therefore, that they were devoid of empirical content and endowed with a universal and fixed character. However, the development of non-Euclidean geometries in the early 19th century showed that other ways of describing space could be discovered and, therefore, that space cannot be *a priori*. Subsequently, the dissolution of the concept of absolute simultaneity and the physicalisation of space entailed by relativity theory seemed to put the nail in the coffin of the universal and fixed Kantian *a priori*. These scientific advancements led most philosophers to discard Kant's epistemological view as fundamentally flawed. Yet, the influence of Kant's view reverberated in the 20th century, for instance, in the work of Duhem and Poincaré, who abandoned the synthetic character of the Kantian 'synthetic *a priori*' and developed the notion of 'convention'. Another example is that of Reichenbach's early works, in which he outlined a developmental Kantianism based on foundational axioms of coordination, with which he aimed at reshaping and modernising the Kantian epistemological system in the light of the radical scientific changes brought about by the non-Euclidean geometries and by relativity theory. These authors named, respectively, 'conventions' and 'axioms of coordination' those propositions that justify the

[7] Cf. Friedman (1999) and Carus (2007) on Carnap's connection with neo-Kantianism. Padovani (2011) and Eberhardt (2011) highlight the neo-Kantian features also of Reichenbach's doctoral dissertation, which he defended in 1915.
[8] Emphasis in the original.

attribution of abstract definitions to empirical items. For conventionalists, this attribution was the result of pragmatic choices, whereas for Reichenbach it involved some deeper meta-theoretical commitment or cognitive function.

The issue of coordination was raised by many empiricists between the end of the 19th and the first decades of the 20th century, particularly Mach, Schlick, and Reichenbach. These scientists and philosophers were concerned with the justification of the ways in which mathematical representations of physical structures, usually presented in the form of equations, could be 'coordinated' with the empirical phenomena that those equations were supposed to describe. In other words, since mathematical structures are themselves devoid of empirical content, how can the theoretical terms which figure as parameters in the equations be given empirical content and thus successfully represent the concrete phenomena? According to Reichenbach (1920/1965), some principles establish the coordination between mathematical representations and their physical correlates, that is, these coordinating principles enable the application of those mathematical tools to empirical reality. For instance, the principle of genidentity (or identity over time) "indicates how physical concepts are to be connected in sequences in order to define 'the same thing remaining identical with itself in time'" (Reichenbach 1920/1965: 55).[9] These axioms of coordination can change along with the development of the mathematical and physical sciences, but retain a *constitutive* character, which was one of the distinctive features of the Kantian synthetic *a priori*. In fact, the meaning of *a priori* which, according to Reichenbach, should be retained, is the one by which we *constitute* the *concept of object*, since the object of scientific knowledge is not immediately given: "Perceptions do not give the object, only the material of which it is constructed. Such constructions are achieved by an act of judgment. The judgment is the synthesis constructing the object from the manifold of the perception" (Reichenbach 1920/1965: 48). In other words, according to Reichenbach, the Kantian notion of 'synthetic *a priori*' can still be relevant to contemporary science, inasmuch as it identifies those preconditions, expressed as principles or functions of thought, that are presupposed by our knowledge claims about the empirical world. These preconditions prescribe norms that enable the representation of our sense-experiences in terms of abstract mathematical concepts.

Reichenbach later abandoned his own interpretation of coordinating principles as constitutive of the object of knowledge after exchanges with Schlick, and talked instead of 'coordinating *definitions*', thus giving up the project of identifying linguistic elements equivalent to the Kantian synthetic *a priori*.[10]

[9] Cf. Padovani (2015).
[10] Cf. Friedman (1994).

The logical empiricist stance on the issue of the status of scientific knowledge converged to a radical separation between a factual element reducible to observation, the 'synthetic *a posteriori*', and a linguistic, 'analytic' *a priori* element, without any room for a synthetic *a priori*.[11] Such a hygienic divide, aimed at preserving the objectivity of science and the realism of the scientific picture, eventually collapsed due to the criticisms coming from two different directions. On the one hand, Quine (1951) famously contested the analytic/synthetic distinction endorsed by the logical empiricists. Instead, he promoted a naturalised holistic epistemology, in which there is no room for any *a priori*.[12] On the other hand, Kuhn (1962/1970) and the following historical and sociological wave in Science Studies challenged the logico-reconstructive character of the bottom-up and observation-first approach of the logical empiricists and argued for the theory-ladenness of observation and the value-ladenness of science. This challenge targeted the logical empiricist assumption of the possibility of synthetic *a posteriori* knowledge coming from purely observational statements and understood the *a priori* in terms of historically shifting pragmatic interests.

In the last decades, philosophy of science has seen the rise of naturalism as its key methodological stance,[13] and an increasing attention to the *practice* of science, as opposed to the reconstructive approach of the logical empiricists. Although most philosophers of science today embrace a well-rounded naturalistic stance, some of them resist *radical* naturalism, i.e., the view that there is no discontinuity between science and philosophy. This resistance often takes the shape of an attempt to redefine a role for the transcendental *a priori* in scientific inquiry, often understood in terms of its constitutive function. In this sense, a line can be traced that – after the logical empiricists – goes from Kuhn's socio-historical interpretation of the *a priori* as paradigms established by communities of inquiry and constituting the 'world' in which they operate;[14] to Hacking's (1992) styles of reasoning as a plurality of scientific methods, which emerge historically and bring about novel ways of constituting scientific entities;[15] to contemporary proposals.

[11] Here, I am offering a highly simplified picture of the early logical empiricist stance on this issue. As it is known, not all logical empiricists held the same view even on this subject. Later in time, some of them defended less radical and more pragmatic alternatives (cf. Carnap 1966).
[12] Cf. below, section 1.1.3 for more on the consequences of Quine's criticism.
[13] Cf. above, section 1.1.1 for a definition of naturalism.
[14] Cf. Hacking (1993) for an analysis on this controversial side of Kuhn's view and for a possible solution. See also Kuhn's (2000) interpretation of his own theory of scientific paradigms as another brand of developmental Kantianism.
[15] According to Hacking (1996: 181), his own styles of reasoning should be interpreted as "a continuation of Kant's project of explaining how objectivity is possible". For a more recent approach, see also Hacking (2015), where he challenges the very possibility of discussing 'objectivity' as a general category.

1.1.3 The analytic/synthetic divide, Quine's holism, and contemporary remarks

In this section, I address Quine's criticism against the analytic/synthetic divide, which was a central tenet of the scientific philosophy of the logical empiricists. Their assumption of a cleavage between a factual and a linguistic component of scientific knowledge was criticised – albeit from different standpoints – by both Quine and Kuhn. The increasing influence exerted by both the naturalistic standpoint in the background of Quine's criticism, and the historicist and sociological turn of science studies prompted by Kuhn and the Kuhnians, eventually led to the collapse of logical empiricism.

As is widely known, Quine (1951) criticised the two entangled 'dogmas' at the basis of the philosophy of the logical empiricists, that is, reductionism and the analytic/synthetic distinction.[16] Reductionism - in this context - is to be understood in the traditional empiricist version of 'epistemic reductionism', in the sense that the set of all synthetic *a posteriori* statements, that is, the propositions of physics, can be derived from one (possibly infinite) set of propositions reporting sense experiences. The dogma of reductionism, according to Quine, is intimately tied to the second dogma. According to the latter, there is a sharp separation between the category of the 'analytic', which refers to the component of a statement that is true in virtue of meaning, and the category of the 'synthetic', which is subject to confirmation or disconfirmation via experience of extra-linguistic facts (for example, sense-perception and experimentation).[17] Quine's criticism is founded on the connection between the two dogmas. As Brittan (1978: 14-15) emphasises:

> [Quine's] basic idea seems to be that those who distinguish between analytic and synthetic sentences do so on the basis of the criterion that synthetic but not analytic sentences have a sense experience translation. But to say that synthetic sentences have a sense experience translation is to subscribe to the dogma of reductionism [since] no two synthetic sentences […] "reduce" to the same set of sense-experience sentences […].

Given that all results in science are fallible, it is always possible to reopen questions, despite their apparent status of settled science. According to Quine, every claim is in principle revisable in the light of new experiential evidence, including the basic rules of logic and mathematics. Therefore, the analytic/synthetic distinction – which, in Quine's view, implies the non-revisability of some principles,

16 In the following analysis, I follow both Quine's own argument and the analysis and critique provided by Brittan (1978).

17 It is worth noting how the logical empiricists' characterisation of the distinction between analytic *a priori* and synthetic *a posteriori* differs from Kant's, despite the frequent misrepresentation of the latter in terms of the former: "Kant does say that analytic judgments are "valid from concepts", but this characterisation is intended to draw our attention to the fact that they are *explicative*, not *ampliative*, and does not commit him to the truth-in-virtue-of-meanings formula" (Brittan 1978: 19, footnote 34).

i.e., the analytic statements – must be dropped altogether. As Quine argues, everything in our total system of knowledge is subject to the 'tribunal of experience' in the very same way, the only difference being in the *degree of justification* between the core and the periphery of this total system.[18] The periphery is where experience mostly impinges and where making changes, that is, replacing components of the system, is more likely and less costly. On the other hand, the core of the system is comprised of highly entrenched statements which, in Quine's view, are shielded from empirical testing, but can ultimately be changed in the light of recalcitrant experiential evidence. Yet, the replacement of very entrenched components leads to many changes in the adjacent peripherical regions. Therefore, changing these components is highly costly from a pragmatic standpoint, because it requires the replacement of all the regions affected by that change. For this reason, we tend to prefer changes at the periphery, rather than at the core of the web-of-knowledge. This is the gist of Quine's epistemological holism, where there is no room for a dichotomy based on two distinct kinds of knowledge, but only on different degrees of entrenchment, which reflect the pragmatic cost of replacing a certain part of the web.

According to a line of argument endorsed, among others, by Putnam (1962, 1976, 1979), Brittan (1978), and Friedman (2000b, 2001), Quine assumes an incorrect equation between analyticity and non-revisability. According to Quine, all statements are in principle revisable. In his view, if a statement is revisable it cannot be purely analytic, i.e., such that it is impossible to confirm or disconfirm it via experience. Therefore, there are no purely analytic statements. However, – these critics claim – non-revisability and analyticity do not coincide. Even if we admit the possibility to revise all statements, we do not rule out that some of them cannot be directly confirmed or disconfirmed from within a given body of knowledge, representational system, or framework of inquiry. Although all parts of a certain scientific framework can be revised, in certain circumstances some of them are not replaced in the light of their empirical disconfirmation but, for instance, because they are no longer suitable to adequately conceptualise an empirical domain, or to provide the required tools for testing a theory or model. In other words, these components are not the best option available to fulfil a certain epistemic *function* within that framework anymore. Therefore, their replacement is not justified in virtue of failed tests against evidence and, in addition, it requires, among other factors, the emergence of alternative conceptual determinations.

[18] The use of 'core' and 'periphery' in Quine is not to be confused with that of Lakatos (1970).

At this point, I must emphasise a radical difference between Quine's understanding of the analytic *a priori* of the logical empiricists and contemporary reappraisals of the transcendental *a priori*.[19] The (analytic) *a priori* that is under the fire of Quine's criticism is characterised as universal apodictic condition for any empirical inquiry, under the assumption of empirical knowledge as a homogeneous whole. Conversely, contemporary views start from the investigation of a context of scientific inquiry and identify the components that can be viewed as constitutive or *a priori* in virtue of their epistemic function within that domain.[20]

In my opinion, there is a direct connection between Quine's fallacy of identifying 'analytic' with 'non-revisable' and his rather simplistic image of a unitary and flat web of knowledge, whose internal changes can all be explained in terms of a single coarse-grained notion, i.e., that of entrenchment. What I suggest is that, if we deconstruct Quine's epistemological holism, it is possible to distinguish between three meanings of it that are left unspecified within his own account:

- Holism (I): with respect to the *unit of acceptance*
- Holism (II): with respect to the *unit of revision*
- Holism (III): with respect to the *epistemic functional homogeneity* of the statements of our theories

According to holism (I), any component of a theory can hardly be incorporated in a body of knowledge (be it a theory, paradigm, or even the whole of science) and considered independently of its relationship with the other components. After the decisive criticisms against the observation-first stance of the logical empiricists, this is hardly contested nowadays. According to holism (II), when conflicting evidence occurs, the revision of statements can be carried out in *any point* of the web-of-belief (or theory, or model, depending on the unit of analysis we take as primitive), and the pragmatic cost of any revision depends on the degree of entrenchment of the statement that we decide to replace. Holism (III), according to Quine, is a direct consequence of (II), since from (II) it follows that there is no fundamental difference in epistemic status between statements. However, even if pragmatic considerations certainly play a role in revising scientific theories and we could intervene, in principle, in any part of the web, it is not actually true that any scientific theory or model meet their experimental test wholesale. As Wimsatt (2007: 103) highlights: "That this thesis is false is demonstrated daily by

[19] As I will discuss in the following sections, the current literature comprises several different stances on whether the components identified as 'constitutive' or 'transcendental' or '*a priori*' should be viewed as analytic or synthetic and in what sense.

[20] Cf. Suarez (2012) for a clear explanation of this aspect in terms of logical quantifiers.

scientists in their labs and studies, who modify experimental designs, models and theories piecemeal […]". Therefore, holism (II) does not lead to holism (III), because a distinction between components with different epistemic statuses can clearly be drawn by looking at their *function within a theory*, which is a central element to be considered in the process of theory-revision, and in scientific change more generally.

What these considerations suggest is that, even if the permanent and cut-off divide between analytic and synthetic propositions is given up, it is still possible and, in my view, desirable, to account for the fact that, given a certain scientific domain, some components have a peculiar status in virtue of their characteristic epistemic *function*, and not simply because they are more entrenched. Their replacement certainly involves considerations of pragmatic cost, but these are not sufficient to account for the 'mechanics' of scientific change, since entrenchment is too coarse-grained as a dependence-relation to capture them, even at a very abstract level. In sum, Quine's holism does not allow us to distinguish between the different epistemic functions of the components within a scientific framework, thus neglecting the importance of these functions when it comes to understanding theory replacement or scientific change in general. As some critics of Quine's holism have argued, his view masks the fact that some components of scientific knowledge perform a specific function of conceptually determining other more empirical elements within a theoretical framework (Friedman 2000b, 2001, Stump 2003, 2015). In conclusion, many contemporary critics of Quine's epistemological holism concede that Quine's criticism against the logical empiricist analytic/synthetic cleavage is successful but argue that it is effective only against a type of *a priori* understood as propositional, devoid of empirical content, non-revisable, universal, and valid for all knowledge. On the contrary, it is ineffective against local, revisable, or non-propositional interpretations of the *a priori*, thus, including most contemporary reappraisals of the constitutive *a priori*.

1.2 The (constitutive) *a priori* in contemporary philosophy of science

1.2.1 Preliminary distinctions

As I mentioned above, dissatisfaction with Quine's holistic replacement of the analytic/synthetic distinction led to the development of various alternatives. Some of these alternatives sought to partly rehabilitate the logical empiricist tradition and to find a way out of its dead end, while others found inspiration in different philosophical traditions, such as operationalism, neo-Kantianism, or pragmatism. In most cases, these alternatives reflect the influence of the historical, sociological, and practice turns in philosophy of science in the second half of the 20th century. As a result, different interpretations of the *a priori* have been suggested in the contemporary literature. According to all of them, the fact that something is to be considered *a priori* as an epistemic precondition for scientific inquiry does not necessarily mean that *a priori* refers to any universal, unchangeable, and necessary proposition. Apart from this shared commitment, the main interpretations that I will present in the rest of this section show major differences that can be clustered as follows:

1. **Unit of analysis of interest.** The issue of the 'analytic *a priori*' was pivotal in the logical empiricist tradition, which focused on the logical reconstruction of scientific *theories*. Therefore, some contemporary interpretations still assume that theories provide the model of what scientific inquiry amounts to. Consequently, they also assume that scientific theories are the unit of analysis of scientific inquiry where (constitutively) *a priori* elements are to be identified. In contrast with this background, other interpretations attempt to translate this issue in terms of scientific *practices* rather than theories.
2. **Character and domain-specificity of the *a priori*.** As I will show, some perspectives understand the *a priori* as completely devoid of empirical content, whereas others allow for it to have some empirical content, even this is not a necessary condition. Another distinction that usually correlates with the former is one between a domain-independent *a priori*, that is, independent of the specific domain of scientific inquiry, as opposed to a domain-specific one.
3. **Relationship between interpretations of the *a priori* and views of scientific inquiry which embed them.** Different interpretations of the (constitutive) *a priori* are usually part of more encompassing views of scientific inquiry, which focus on some units of analysis rather than others (e.g., theories, paradigms, etc.), provide competing explanations of scientific change, and make assumptions concerning the relationships among the epistemic dimensions of scientific inquiry (theorising, experimentation, measurement, modelling, etc.). Therefore, each specific interpretation of the *a priori* is tied to – and, in some cases, influences – features of the more general view of scientific inquiry in which it is embedded. When comparing different interpretations, I will pay attention both to their implications with respect to other philosophical issues about science, and to whether it is more appropriate to understand different interpretations as alternative to one another, or as compatible.

1.2.2 Friedman's relativised *a priori*

In *Dynamics of Reason* (2001), Friedman offers his interpretation of constitutive *a priori* elements as changing relative to the growth of science. His interpretation is the most influential and one of the best-developed among recent reappraisals of the *a priori* in scientific inquiry, and it is embedded in Friedman's larger project of tracing a role for philosophy in the development of science from a Kantian standpoint.

Friedman endorses Kuhn's (1962/1970) view of scientific change, based on the distinction between periods of normal science and of revolutionary science. During normal science, a scientific paradigm is *constitutive* of the daily problem-solving activities of scientific inquiry. A paradigm is – according to Kuhn – a single set of generally agreed-upon rules of inquiry, which is established after a pre-paradigm state of conflicting schools and defines the normal state of scientific investigation prior to a scientific revolution.[21] Friedman interprets Kuhnian paradigms as linguistic frameworks in Carnap's sense. According to Carnap (1950), in the different linguistic frameworks or formal languages within which we can formulate our scientific theories, different logical rules or principles are constitutive of the concepts of 'validity' and 'correctness' of the statements of our theories, therefore determining important characteristics of the mathematics, logic, and ontology of a theory.[22] For example, we could choose a framework where intuitionist logic holds, rather than classical logic, or a framework that takes sense data as primitives or physical objects as primitives. Analytic *a priori* statements, in Carnap's view, define the rules constituting the frameworks, while synthetic *a posteriori* statements express the empirical laws that result from observational data, as we saw in the previous section. However, the truth, correctness, and validity of the empirical statements cannot be evaluated without reference to the standards dictated by the framework itself. The Carnapian theory of linguistic frameworks is employed by Friedman as a tool to give a more formal interpretation of Kuhnian paradigms which, according to Friedman, are composed of *relativised a priori* principles.[23]

[21] Famously, Kuhn (1970) himself gave several definitions of a scientific paradigm (cf. below, footnote 23). Here, I only consider the one central to Friedman's account.
[22] For more on Carnap's view of the constitutive role of the linguistic frameworks, see Parrini (2009).
[23] For a criticism of Friedman's interpretation of Kuhnian paradigms as informally formulated Carnapian frameworks, see Mormann (2012). Mormann reminds us of the many different and, somehow, elusive definitions of 'paradigm' provided by Kuhn, as pointed out by Masterman (1970), and contrasts it with the interpretation of Kuhnian paradigms as frameworks composed of relativised *a priori* elements already endorsed by Friedman (1993). For another excellent criticism of Friedman's interpretation of Kuhn's view, see Richardson (2002).

By focusing chiefly on space-time theories, Friedman argues that physical theoretical frameworks are comprised of a mathematical, a mechanical, and an empirical component, where the mechanical part provides the *coordination* – that is, it establishes and justifies the referential relationship – between the mathematical and the empirical parts.[24] For instance, in Newtonian mechanics, the three laws of motion are the mechanical part, in that they establish the coordination between the mathematical part, supplied by the mathematics of the infinitesimal calculus, and the empirical part, which is the domain of concrete phenomena described by the law of gravitation. Within the context of a physical theory, the mathematical part and the mechanical part are together *constitutive* of the empirical part. The mathematical and the mechanical parts are, therefore, *a priori* but only in the sense that they are constitutive of the fundamental theoretical concepts of a theory. According to Friedman, this means that they provide elements of language and the conditions of testability for the empirical statements. In other words, they enable the statements expressing concrete physical regularities to have empirical significance and, therefore, to have a truth-value.

These constitutive principles are not fixed but change during profound conceptual revolutions that bring about a novel paradigm. According to Friedman, the transition from Newtonian mechanics to relativistic physics was marked by a major conceptual shift involving a change in the paradigm's constitutive principles. The paradigm of classical mechanics framed by three-dimensional Euclidean space and the Newtonian laws of motion was replaced by that of special relativity, where the four-dimensional geometry of Minkowski space-time and the light principle are constitutively *a priori*, and by that of general relativity, where the geometry of (semi-)Riemannian space-time manifolds and the principle of equivalence are constitutively *a priori*. In the words of Mormann (2012: 29), Friedman sees the "evolution of scientific reason – encapsulated in a sequence of theories or conceptual frameworks – as a continuous trajectory in a logical or conceptual space".

It is crucial to highlight that Friedman's *relativised a priori* strictly belongs to the domain of mathematical physics, for two main reasons. Firstly, this aspect is salient to his argument against Quine's epistemological holism. In Friedman's view (2001: 80), the mathematical part of our theories is not itself tested directly against empirical evidence, since "[…] what is empirically tested is rather the particular coordination in virtue of which some or another mathematical structure is used to formulate precise empirical laws about some or another empirical phenomena". In the second place, the *relativised a priori* are contrasted by Friedman with a more general notion of presupposition as constitutive

[24] By identifying the chief role of the constitutive *a priori* as means to solve the coordination problem, Friedman's view explicitly builds on that of Reichenbach (1920/1965). Cf. above, section 1.1.2.

condition for the truth of an empirical statement. According to this notion, all empirical statements in principle include implicit presuppositions, which must be assumed as true so that each empirical statement can have a truth value. For example, the statement 'Carbon-based molecules have three general types of structures' lacks a truth value, if the presupposition that 'there exist carbon-based molecules' does not hold. Friedman's *relativised a priori* are not presuppositions of this general sort. The *universal* character of his constitutive principles lies in the fact that they provide the structure of the constitutive spatio-temporal frameworks within which *all* empirical inquiries take place, at a certain stage of scientific development: "We want to reserve this characterization for *particularly fundamental presuppositions* lying at the basis of mathematical physics – principles which, accordingly, can plausibly be taken as fundamental presuppositions of all empirical truth" (Friedman 2001: 74).[25] However, his relativised *a priori* principles are not universal in the sense criticised by Quine, in virtue of the fact that they are involved in any scientific knowledge. Yet, they are universal because they are involved in that domain of scientific knowledge which, according to Friedman, provides the basis for all empirical inquiry.

1.2.3 Alternative contemporary interpretations of the *a priori*

1.2.3.1 Stump's functional and contextual a priori

Stump (2015) develops an account of constitutive elements in science that he considers alternative to Friedman's. The main motivation behind his interpretation is to reconcile a role for the constitutive *a priori* with the fallibilist maxim, according to which all results in science are fallible. Stump's view sets off from the fact that some knowledge must be already in place for us to be able to implement a scientific inquiry in a specific context. However, even those elements that were deemed as settled once and for all can change. Yet, this fact should not – according to Stump – hasten a straightforward holistic perspective, apparently more in harmony with fallibilism. Clearly, there are some theoretical elements that do not change simply because of an accumulation of empirical evidence that disconfirms them, but because they are no longer adequate to categorise the empirical data that they are supposed to account for. The fact that there are some elements in a theory that should be deemed as constitutive does not mean that these elements have to be devoid of empirical content, as already emphasised by Friedman. On the one hand, this consideration leads Stump to view the label '*a priori*' as outdated, given the dynamic nature of these components. On the other hand, it shifts the focus of the epistemological analysis towards the *function* of these elements, which is overlooked by holistic

[25] Emphasis mine.

accounts: "These principles are not only 'hardened' so that no one would seriously doubt them, but also play a special role in categorizing phenomena" (Stump 2015: 4).

The notion of 'constitutive' captures this peculiar character of certain theoretical principles, which differ from other more empirical statements of a theory but not due to a difference in their degree of entrenchment, as already argued by Friedman (2000b, 2001). However, Stump's view departs from Friedman's relativised *a priori*, in that the distinction between constitutive and non-constitutive is not permanent. According to Stump, who builds on the notion of the functional/contextual *a priori* defended by Pap (1944), the constitutive function of certain elements is not *absolute*, but can change depending on time and context. Not only are there no necessary and sufficient conditions for something to be constitutive, we do not even need them, since the constitutive role is determined by the context, not by their essence. There is no way to identify specific statements as constitutive before going into the details of specific scientific inquiries. Yet, the contextual distinction between constitutive and non-constitutive is very useful for a more accurate account of how science works: "If you want a nuanced and accurate account of science, then you must consider what is a precondition for something else, what roles various claims are playing at a given time and context, and how these change over time" (Stump 2015: 16).

Stump's account rules out the possibility of a search for *universal* principles, as envisioned by Friedman, that is, principles that are valid for all science at a certain stage of development. In other words, Stump relaxes Friedman's strict focus on fundamental mathematical-physical presuppositions to extend the label 'constitutive' to further presuppositions and he does not specify whether there are any criteria relative to the fundamentality of these presuppositions. In this sense, Stump goes towards a view of constitutive elements better apt to describe a larger variety of epistemic domains. However, he stays close to Friedman's domain-specific interpretation of the constitutive *a priori* and to the focus on theories only, whereby constitutive elements are internal components of theories. In addition, he does not offer a systematic view of *how* context and history can make room for a more flexible notion of 'constitutive' in more detail.

1.2.3.2 The a priori in epistemology of measurement and thought experimentation

As I mentioned above, Friedman's view of the relativised *a priori* principles is inspired by Reichenbach's axioms of coordination. According to Friedman's interpretation of Reichenbach's work, certain principles internal to a physical theory provide the coordination between mathematical representations and physical correlates. In other words, these principles enable the representation of empirical reality by means of mathematical tools. Friedman's work stimulated further research on how coordination is justified – within different theories – by certain theoretical principles, such as the light principle, the principle of equivalence, and the principle of least action (e.g., Ryckman 2005, Stöltzner 2009). According to these philosophers, such principles have "a fundamental structural significance for the theory" (Padovani 2015: 123), that is, they have a fundamental role as components *internal* to the theoretical structure itself. These principles must be viewed as constitutive because they provide fundamental elements of the conceptual framework within which the theory can be formulated and empirically tested.

In his recent but influential work on scientific representation, van Fraassen (2008) tackles the debate on coordination from another angle, more specifically, with relation to the role of measurement in scientific inquiry. Our measurement practices presuppose certain classifications provided by a theory of reference with respect to, for instance, the salient parameters to be measured and the relationships between them. Yet, how can we trust a theory before we have empirical support, which is usually obtained via measurement? Given this problem, according to van Fraassen (2008: 121) the function of coordination is to "determine how measurement can establish a value for what is measured". Therefore, van Fraassen criticises Reichenbach's analysis of coordination and, thus, implicitly also Friedman's reinterpretation. According to van Fraassen, the axioms of coordination alone cannot account for the *historical* process through which a certain measured parameter and the procedure to measure it are coordinated. How could *theoretical principles* provide the conditions to relate mathematical structures to empirical phenomena, especially if considered a-historically, in the abstract, as van Fraassen interprets Reichenbach's attempt?[26] According to van Fraassen, we cannot accomplish this task by abstracting away from the coordination already in place when a new one is established. A form of coordination cannot be achieved in isolation from its historically prior form, and the concurrent historical development of theory and measurement procedures must be analysed to understand how a new form of coordination is established.

I will discuss the specific issue of coordination in the context of contemporary epistemology of measurement in chapter 4. Here, I only want to address how even a well-rounded empiricism that leaves no room for any *a priori* in a traditional sense, such as van Fraassen's, still seems to imply that something about our ways of knowing contributes to justifying the representational relationship established in the coordination. According to van Fraassen, scientific representation cannot be

[26] Padovani (2017) draws attention to the elements of practice and measurement considered by Reichenbach himself, especially in his early work, in contrast with van Fraassen's interpretation.

considered as a mere mimetic enterprise. Scientists trade on selective likeness, unlikeness, distortion, addition, and abstraction of aspects of a target system to represent it for some purpose. Scientific representation, in his view, does not encompass theorising only, but also experimentation and measurement. Considering measurement as a form of representation means that what measurement shows is not directly what the measured is like *per se*, but how it appears in a specific measurement set-up. It is in this sense that van Fraassen suggests that quantities are not to be considered as properties of an object of inquiry independently of our ways of representing it. Interpreting certain measurement outcomes as values of a property requires a form of coordination between what is represented and its representation that is, according to van Fraassen, justified by the theory governing the experimental and measuring context of inquiry. This way of understanding how the measured parameter comes to acquire meaning through coordination may be considered as an empiricist interpretation of the constitutive *a priori* (although hardly *a priori* at all, in this case), since it assumes that certain conditions, albeit empirical and historically changing, provide justification for the coordination between representation and represented. In addition, further work on this issue suggests that it is not just theory that provides justification for coordination in measurement, but certain material conditions of inquiry can sometimes play a justificatory role (e.g., Baird 2004, Padovani 2015, Tal 2017a).

Another debate shows that there are interpretations of the (constitutive) *a priori* alternative to its understanding in terms of internal components of theories and involving epistemic dimensions other than pure theorising. I am referring to part of the literature on scientific thought experiments (TEs) and their relationship with real experiments (REs). Whereas classic views in this context were developed with reference to Platonist approaches (Brown 1991, 2004) or empiricist ones (Norton 1991, 2004), alternative stances use different interpretations of the transcendental *a priori*. More specifically, Buzzoni (2008, 2013a, 2013b, 2018) characterises empirical thought experiments as counterfactual anticipations of experimental apparatuses in terms of their conceptual content. Thought and real experiments do not differ from the empirical-operational point of view, since both are conceived in the light of a theory and a well-specified experimental apparatus. In fact, the specific content of any empirical thought experiment *must* be, at least in principle, reducible to empirical-operational interventions on reality, otherwise it would lose its empirical character (Buzzoni 2013b). The difference between thought and real experiments lies only on a transcendental level, since thought experiments often function as general preconditions of real experiments:

> TEs are the conditions of possibility of REs because, without the a priori capacity of the mind to reason counterfactually, we could not devise any hypothesis and would be unable to plan the

corresponding RE that should test it. This capacity underpins the distinction in principle – a properly transcendental distinction – between TEs and REs (Buzzoni 2018: 333).

Buzzoni's account of thought experiments is founded on his understanding of the transcendental *a priori* as devoid of any empirical content, the latter being given through sensations only. More precisely, Buzzoni interprets the *a priori* as "the capacity of the mind to consider things counterfactually" (Buzzoni 2018: 336), rather than as a set of propositions. Performing thought experiments is possible in virtue of this capacity of the mind, which allows us to manipulate ideal entities in an imagined experimental setting. It is in this sense that thought experiments are conditions of possibility of real experiments. In so far as a thought experiment presents a specific counterfactual situation, whose empirical details are reducible to that of the corresponding real experiment, it instantiates the capacity of the mind to reason counterfactually.

Buzzoni's interpretation of the transcendental *a priori* and of scientific thought experiments is rejected by Fehige (2012, 2013), who contrasts Buzzoni's understanding of the *a priori* with Friedman's interpretation. According to Fehige, thought experiments depend on a theory of reference and, therefore, are justified by the constitutive *a priori* principles internal to this theory of reference. A third interpretation is introduced by Stuart (2017), who suggests that thought experiments can fix the meaning of different theoretical structures (laws, equations, models, concepts, etc.) by using imaginary examples. Stuart's interpretation of the *a priori* envisions a more dynamic understanding of the relationship between theorising and thought experimenting in terms of, again, the reasoning *ability* that allows us to establish connections between concepts and between concepts and experience.

In conclusion, I want to spell out what these discussions of measurement and thought experiments tell us about the contemporary interpretations of the (constitutive) *a priori*:

- Both debates emphasise how experimentation, theorising, measurement, and other epistemic dimensions are often highly interconnected. Conceiving the role of the (constitutive) *a priori*, in any of its interpretations, as limited only to certain cases of highly abstract theorising in space-time physics seems too restrictive.
- Both debates show that there is variability with respect to how contemporary interpretations are tied to traditional features of the *a priori*, such as necessity, lack of empirical content, universality, and fixity.
- Developments in the debate on measurement emphasise that certain constitutive conditions of inquiry may be identified along the material dimension of experimentation, in contrast to the conception that constitutive *a priori* elements are only internal components of theories.

In the next section, I introduce one last view, which develops an interpretation of the transcendental *a priori* that is consonant with a practice-oriented perspective on science.

1.2.3.3 Chang's practice-oriented interpretation of the a priori

Building on the work on the *a priori* by C. I. Lewis,[27] Chang (2008, 2009) uses a transcendental approach to identify some principles that identify the epistemic preconditions for science as a set of *practices*, rather than as a body of propositional knowledge. Chang conceives these principles as something we need to assume in order to perform scientific inquiries, in contrast with interpretations of (constitutively) *a priori* principles as embedded in specific scientific theories or theoretical frameworks. Chang explicates a set of basic principles that must be held as valid, depending on the different epistemic *activities* that an individual or community aims to implement. He defines an epistemic activity as "a coherent set of mental or physical actions (or operations) that are intended to contribute to the production or improvement of knowledge in a particular way, in accordance with some discernible rules (though the rules may be unarticulated)" (Chang, 2011b: 209).[28] One example of basic epistemic activity is that of testing by overdetermination. This activity can consist in checking the prediction of a theory by comparing the predicted value to the observed value, but it can also be a comparison between two (or more) theoretical determinations or two observational ones. If we want to engage in this activity, according to Chang, we must assume the *principle of single value*, according to which a real physical property cannot be ascribed more than one definite value in one and the same situation. Another example of basic epistemic activity is that of counting. To perform this activity, we must assume that the domain of interest is composed of discrete elements, that is, we must assume the *principle of discreteness*. Otherwise, the activity of counting would not make sense in that domain.

These principles are neither empirical generalisations, nor logical truths, and each is required to carry out a specific basic epistemic activity. They certainly cannot be considered as *universal* principles in Friedman's sense that they are tied to fundamental theoretical frameworks at a certain stage of scientific development. Yet, they are foundational for certain epistemic activities that the epistemic agents want to perform. Scientists choose contingently and pragmatically the epistemic activities they

[27] Cf. especially Lewis (1929).

[28] Cf. also Chang (2014: 71-72). Chang's broader framework based on systems of practice and epistemic activities is outlined in Chang (2011b; 2012, chapter 5; and 2014). His view on epistemic objects, as bound up with the systems of knowledge to which they belong, is expressed in Chang (2011a), whereas he develops his notion of pragmatist coherence in Chang (2017).

want to implement, and this determines which principle(s) must be held as valid in a context. This interpretation of the *a priori* has two main implications:

- We see a shift of focus from truth to intelligibility. These principles make our epistemic activities *intelligible*, rather than making our theories true or empirically confirmable, and, in this sense, they determine what kind of features the entities we investigate may have. For example, in the case of testing by overdetermination, if we do not assume the principle of single value, the activity will seem unintelligible. The principle itself determines an essential feature of the *possible object* of investigation, independently of the specific domain of entities considered, in the sense that *any* physical property cannot have two different values in the very same situation.
- Changes in the use of these principles are determined by purely pragmatic criteria that depend on the *choice* of the epistemic subject(s) to carry out a specific epistemic activity for some purpose. There is no restriction on the variety of epistemic activities that different epistemic communities might want to engage with in various times and places. The normative character of these principles lies in their stability and permanence, which results from their formal character.

This last point best illustrates the difference with Friedman's interpretation of the constitutive *a priori*. On the one hand, Chang does not identify any (constitutively) *a priori* principles within specific theories. By shifting the focus to practice, the unit of interest becomes that of the *activities* performed by scientists, rather than the theories they produce as an outcome of their activities:

> A serious study of scientific practice must be concerned with what it is that we actually do in scientific work. This requires a change of focus from propositions to activities. I begin with the recognition that all scientific work, including pure theorizing, consists of actions—physical, mental, and 'paper-and-pencil' operation. (Chang 2011b: 208).

As the quote shows, Chang subsumes theories under practice, by understanding them as a temporary product of a process composed of complex epistemic activities, more fundamental, in a sense, than the theories it produces. This aspect and the formal (i.e., contentless) character of his principles are the two basic differences between Chang's view and Friedman's. Yet, their interpretations of the *a priori* are not mutually exclusive, but can be looked at as compatible with one another. While Friedman's relativised *a priori* principles are theory-relative, that is, they change along the development of science, since they can have empirical content, Chang's principles are located at a meta-level when compared to Friedman's, in that they are not just domain-independent, they are not about theory *at* *all*. Still, Chang's broader view crucially diverges from Friedman's in that the former does not allow for any universal or fundamental set of principles at a time while, in the latter, the constitutive role is played only by particularly fundamental principles of mathematical physics, universally valid at a certain stage of scientific development. In addition, Chang's principles do not change over time because of radically revolutionary changes in fundamental scientific frameworks.[29] The principles, rather, are chosen and replaced on purpose, and the flexibility and freedom in the *use* of these principles by any epistemic community are closely tied to Chang's particular conception of epistemic subjects.

It is worth emphasising that a pragmatic view of science is in the background of Chang's interpretation of the transcendental *a priori*. In discontinuity with the traditional abstract philosophical understanding of notions like truth, explanation, confirmation, or reduction, Chang (2011b, 2012, 2014) aims at reintroducing both the notion of epistemic subject as agent and the centrality of practice in the philosophical analysis of science. On the one hand, his stance mirrors the practice turn in philosophy of science, initiated by philosophers who strove to provide an analysis of scientific activity more consonant with experimental practice.[30] On the other, it gives attention to the theme of embodied knowledge and know-how that dates back to Polanyi and Gooding.[31] Given this background, instead of identifying (constitutively) *a priori* statements that justify the representation of empirical phenomena via mathematical tools, Chang focuses on preconditions that underlie the possibility to perform certain epistemic activities required to pursue scientific inquiries in general.

1.3 Motivation and methodological preliminaries

1.3.1 Why bother?

So far, I have introduced different contemporary interpretations of the (constitutive) *a priori*. All of them represent epistemological stances alternative to Quine's holism, because they suggest that some components involved in our scientific inquiries stand out, as far as their epistemic function is concerned, as enabling conditions for empirical knowledge, rather than because they are simply more entrenched. Yet, I have emphasised the great variability, among these interpretations, with respect to the characterisation of these elements in the context of contemporary science. First, some

[29] As I mentioned, Chang reinterprets Lewis' pragmatic *a priori*, which does not allow for a relativisation in the sense put forward by Friedman and the Reichenbachian tradition. It is the changeability of the principles, according to the epistemic purpose tackled by each epistemic community and, therefore, to a context of inquiry, that characterises them as flexible and not universally fixed. Cf. Chang (2011b: 214).
[30] Cf., among many others, Hacking (1983) and Rheinberger (1997).
[31] Cf,, for instance, Polanyi (1966) and Gooding (1990).

philosophers, in continuity with the logical empiricist tradition, identify these elements with individual statements, whereas others view them as general functions of thought. Secondly, some characterise these elements as internal structural components of theories and, therefore, as domain-specific, while some others identify components that are not specific to any scientific inquiry. Finally, interpretations of the *a priori* whereby it can have empirical content are contrasted with contentless interpretations.

Depending on the interpretation of the (constitutive) *a priori*, we can identify how these perspectives trade on certain features of *a priori* elements as they were conceived in the rationalist tradition, that is, as necessary, universal, and fixed propositional elements. As I showed in paragraph 1.1.2, the Kantian view that some components can extend our knowledge of the world while, at the same time, being in some sense *a priori* was highly contested and eventually discarded exactly because it was interpreted as the claim that certain propositions of some sciences, Euclidean geometry and Newtonian physics in particular, were *a priori* in the standard rationalist sense of this expression. For something to be *a priori*, according to critics of the Kantian perspective, it had to be an unchangeable, necessary, and universally true proposition. This interpretation certainly influenced the unfolding of the entire analytic/synthetic debate of the 20th century. Therefore, the views of the *a priori* that I introduced in the previous section can be considered as attempts to redefine a notion that may still be useful in contemporary analyses of science by disconnecting it from certain features it was wrongfully attributed to in the first place.

Friedman's move, in the light of his interpretation of Reichenbach, is to preserve the propositional character of the *a priori*, while relativizing its fixity to scientific paradigms or theories, so that the relativised *a priori* elements change when a scientific paradigm is overturned and replaced by a new one. Therefore, the necessary character of the constitutively *a priori* principles can be grasped only in the light of the historical process that led to the formulation of new constitutive spatio-temporal theories, within which they played a peculiar epistemic function. In this sense, Friedman's principles are necessary only in a local (i.e., within a framework) and historical (i.e., as part of a succession) sense. The universality of these *a priori* components is lost in the historical dimension, but is preserved synchronically, in that the set of relativised *a priori* principles that Friedman identifies is supposedly valid for all science at a certain stage of development. Stump pushes the relativisation of necessity, fixity, and universality even further. According to his view, these components are fixed and necessary only relative to a certain epistemic function they perform within a framework, while universality is replaced by the importance of context to identify their function. Although he drops their characterisation as *a priori* altogether to preserve only their 'constitutive' label, he does not question the exclusively propositional character of these components. Buzzoni follows quite another path, in that his *a priori* is not propositional at all, but rather identifies a way of the functioning of our mind, by which we manipulate ideal entities counterfactually. This capacity is universally valid and absolute, in the sense that it is not legitimised, or justified, by other epistemic components. Yet another move is made by Chang, who preserves the fixity of his principles, since they do not necessarily change in character across time, and their universality, because they can potentially be deployed in any scientific framework. What can change is the context in which and the epistemic goal for which they are used, while their necessary character is of a *pragmatic* kind, since it is relative to the pragmatic choices of the epistemic subject(s). Although his principles are still propositional in character, they are required to perform certain epistemic activities, rather than to justify abstract representations of concrete phenomena.

At this point, sceptics of non-holistic views may ask two questions. Firstly, whether it still makes sense to characterise something as *a priori* in scientific inquiry in an epistemological sense. As we have seen, contemporary interpretations dismiss the naturalist argument by which the development of science has shown that some propositions thought to be *a priori* were eventually rejected due to new empirical evidence. This characterisation of the *a priori* simply has no place in scientific inquiry. Yet, sceptics might emphasise that, once we drop traits such as the lack of empirical content, necessity, universality, fixity, or propositional character, of supposedly (constitutive) *a priori* elements we end up having something very different, if at all similar, to what is traditionally conceived as *a priori*. If we have to drop the idea that any part of our knowledge system is *a priori* in the traditional sense, we might as well just think that, perhaps, only certain social conditions of scientific inquiry, or certain values, or perhaps nothing at all has something like an *a priori* character, but certainly nothing in an epistemologically salient sense. Secondly, they may ask what the reason to identify those elements would be, especially considering that contemporary interpretations seem to refer to heterogeneous kinds of epistemic components.

My answer to both questions develops one of the few common traits shared by all contemporary interpretations of the *a priori*. What all these views share is that they identify as (constitutively) *a priori* those elements that have a specific *function* of justifying, that is, of legitimising, either a certain mode of production of empirical knowledge, or a certain outcome of scientific inquiry, rather than the fact that they have a certain property or set of properties (lack of empirical content, necessity, universality, etc.) in a binary sense. To the question "How is our scientific knowledge valid?", we cannot simply answer, according to these views, "Because the world is such and such", but we must answer "Because, if we attribute this epistemic function to this feature of our knowledge system, the world is such and such". Therefore, the motivation to propose yet another characterisation of the *a priori* is to continue the debate on how scientists find justification for the empirical knowledge of the world that they obtain, and for the ways in which they obtain it. This justification essentially involves the interactions between the epistemic subjects or communities, their cognitive abilities, the epistemic tools they use, and the environment that surrounds them. Crucially, justification is not a once-and-for-good business but is itself an achievement, whose form and content can change even quite radically across time. Therefore, it is crucial to investigate it as a situated historical enterprise in which contextual factors matter, when it comes to identifying those conditions that enable its establishment.

In the rest of this dissertation, I will focus on this issue of justification with respect to the problem of coordination, the history of which I briefly introduced in section 1.1.2, as it is variously discussed in contemporary philosophy of science. The problem of coordination, in its most general form, refers to the struggle that scientists meet in order to justify their representation of concrete phenomena in the real world by means of abstract representational tools, which, by themselves, do not refer to anything empirical. My specific goal will be that of identifying, within specific and historically situated scientific inquiries, those epistemic components which contributed to the justification of a certain form of coordination. In other words, I will individuate those elements that functioned as required epistemic (pre)conditions for a form of coordination to be established and justified. This is the fundamental feature behind my understanding of the constitutive *a priori*.[32]

Since I emphasise the epistemic function of these components over their possession of certain specific features, I follow Stump and drop the label of *a priori*, which historically carries the requirement of being devoid of empirical content, but retain the label 'constitutive', which indicates their epistemic role of providing part of the justification for the referential relationship between abstract representational tools and concrete phenomena. Yet, in the next chapter, I will identify some typical features shared by these components. However, these traits admit of degrees and must be conceived as tools to identify the constitutive components in the context of a form of coordination, rather than as binary defining properties. By means of these analytic features and the case studies that I examine, I will eventually suggest that there are qualitatively distinct kinds of components involved in justifying forms of coordination, and that they may be referred to as 'constitutive' to different degrees.

[32] I provide a more precise definition in chapter 2, section 2.1.

1.3.2 History, practice, and the heritage of theory-ladenness

As I have explained, the general notion of 'coordination' refers to the justification for the use of abstract representational tools, that is, of theoretical concepts and symbolic expressions, to refer to real empirical phenomena. Therefore, using the notion of 'constitutive' to indicate those components which contribute to establishing and justifying a form of coordination requires outlining the relationship between this notion and the broader discussion on the role and character of scientific concepts. Conceptual change in science can follow several patterns, and many kinds of causes can concur in prompting or pre-empting it, at both the individual and collective level (Love 2015; Nersessian 1992; Soler et al. 2008; Thagard 1992, 2012). In addition, contemporary philosophical accounts of scientific concepts characterise them in different ways according to their epistemic role, from the outcome of taxonomic practices (Hacking 1996, Kendig 2015, Kuhn 1990), to research tools for investigating phenomena (Feest 2010, Feest & Steinle 2012, Steinle 2016), to solutions constructed to answer specific problems (Nersessian 2010).

In the next chapter, I will discuss more in detail my specific understanding of constitutive elements and their relationship with the problem of coordination between abstract representational tools and empirical phenomena. Here, I just want to spell out some methodological precepts that will guide the rest of my work. Firstly, all branches of the contemporary literature on scientific concepts embrace a pluralistic approach, whereby the attempt to reduce scientific concepts to a single homogeneous class, either based on their role, of some of their properties, or of their genealogies, hardly makes any sense. In the rest of this dissertation, I will retain a pluralistic outlook on the variety of conceptual tools, be them laws, quantity concepts, or mathematical models, the coordination of which with concrete phenomena will be the subject of my analysis. In short, for the purposes of this work, I will not assume that certain classes of conceptual tools are in principle reducible to others. Secondly, I want to emphasise the fundamental importance of the historical dimension for the analysis of coordination and of constitutive elements in scientific inquiry. In this sense, history of science is not a mere source of case studies, but it essentially enables the individuation of those factors involved in how referential relationships change across time, together with the ways in which they are justified (Curiel 2018). Finally, my analysis of how forms of coordination are achieved and what role constitutive elements play in their justification will focus also on how justification is obtained *in practice*. This means that I will not consider coordination as something ready-made, but as the result of an active process, something that scientists accomplish and then preserve or modify along the development of science. Yet, someone may object that, in order to reconstruct how coordination is justified, we do not need to look at practice, that is, at the process of achieving that justification, but it is enough, and simpler, to focus on the final product. Clearly, my approach is descriptive and reconstructive, since I do not aim at making normative claims concerning the process through which scientists achieve and justify coordination, but merely at describing its historical unfolding and epistemic grounds. Yet, the focus on practice, or, as some would say, on the context of discovery – i.e., on the process by which coordination is obtained and not only on coordination simply as a result – is not at all irrelevant to reconstructive approaches. Considering *how* scientists came to consider a certain 'product' as a stable and justified form of coordination is a fundamental source of insight. In fact, it enables us to better understand which preconditions had a crucial epistemic role in the achievement of coordination and, sometimes, it is helpful to dispel mistaken assumptions concerning the character of coordination itself, especially with respect to historical episodes rather far in the past.

An issue in the background of the discussion of coordination and constitutive elements regards the units that a philosophical analysis of science and scientific change should adopt. As we have seen, the logical empiricist tradition viewed theories as the privileged unit of analysis, and this influenced Friedman's (2001) interpretation of Kuhnian paradigms – his chosen unit for analysing coordination and his constitutive *a priori* principles – with a strong bias towards theory. Since this dissertation also focuses on the coordination between abstract conceptual tools and empirical phenomena, it may seem straightforward that the proper unit of analysis to adopt should be scientific theories, since it is in theories that concepts, laws, and principles figure. Yet, theorising encompasses several basic epistemic activities such as classifying, calculating, comparing, hypothesising, etc. In addition, it is often highly entangled with other epistemic dimensions, including experimentation, measurement, etc. The category of 'theory' itself has shifting boundaries, depending on the discipline or specific scientific inquiry and its degree of maturity. In the rest of this section and in the following, I will discuss a notion that has been quite influential in philosophy of science in general and in the debate on the constitutive *a priori* specifically, i.e., theory-ladenness. Understanding the influence of this notion is a key methodological point, because it is essential to free my analysis of coordination and constitutive elements from the two assumptions that theories are the only relevant unit of analysis, and that theorising is the only dimension along which constitutive elements can be found.

Many epistemic components of scientific inquiry along the epistemic dimensions of observation, experimentation, measurement, and scientific instrumentation are often said to be theory-laden. The general idea behind theory-ladenness is that our access to the world (in the form of data, phenomena, etc.) is always mediated by our perceptual modalities and/or by our conceptual resources. Therefore, theory-ladenness seems to be primarily about our observations and, derivatively, about all those activities and tools which are supposed to give us direct access to observational data (e.g., experimentation, measurement, instruments). In what follows, I briefly reconstruct the historical development of this notion to show its shifting and increasingly vague meaning.

The strongest version of theory-ladenness is sometimes attributed to a stereotypical logical empiricist standpoint, with Popper often addressed as one of its main supporters.[33] According to this view, all experiments are (or should be) explicit tests of existing theories about the objects in question. Yet, the logical empiricist tradition viewed experiments as the privileged way to test scientific theories thanks to their direct access to phenomena via observation. Therefore, such a view does not seem to qualify as theory-ladenness in the way it was subsequently characterised by scholars reacting to logical empiricism. Many opponents of logical empiricism argued for the theory-ladenness of observation, that is, the fact that our experience is organised through categories already determined by our theoretical stances. According to this view, since experimentation requires observation, and observation is theory-laden, also experimentation is, therefore, theory-laden. According to Hanson (1958), any causal claim made via observational statements is theory-laden, since any attribution of causal connection is made in the light of a theoretical background. Kuhn (1962/1970) famously develops the notion of theory-ladenness in terms of paradigm-ladenness: Scientists working from within a scientific paradigm are committed, among other things, to certain shared theoretical assumptions that determine their ways of engaging with their objects of inquiry including, to an extent, their modes of perception. The Kuhnian influence was predominant for the two decades after the publication of *The Structure of Scientific Revolutions*, leading to various forms of scientific anti-realism founded on the theory-ladenness of observation and instrumentation (e.g., Kosso 1989). Its long echo reaches the current debate with van Fraassen (2008) who, as we have seen, emphasises the constructed and theory-laden character of scientific observation through measuring instruments.

Starting from the late 1970s, an alternative standpoint on the relationship between theory, experimentation, and observation started to emerge. Most famously, Hacking (1983) argued against the theory-ladenness of experimentation. Hacking claimed that Newton's experimental observations of the dispersion of light preceded any theoretical interpretation. According to him, scientific experiments can give us direct access to phenomena via manipulation, thus justifying a form of experimental realism in contrast with the anti-realism resulting from post-empiricist theory-ladenness. After Hacking, the possibility of theory-free experiments is usually defended by appealing to the fact

[33] Cf., for instance, Chalmers (2003: 493) and Radder (2003: 161).

that experimentation can be conducted in the absence of theoretical knowledge, but with knowledge of the causal power of the instrumentation used. Within this camp falls Steinle's (1998: 284-292) analysis of Roentgen's experiments on X-rays – previously examined by Kuhn (1970: 57) – in the light of the category of 'exploratory experiments'. These experiments are "not conducted to test a theory, but to expand our knowledge of causal connections in relation to the scientific instruments and devices involved" (Heidelberger 2003: 144). In addition, claims of theory-free experiments often receive support by various historical analyses of how a certain scientific apparatus worked without agreement on exactly how it did so from a theoretical point of view.[34] Finally, appeal is also made to the fact that the use of scientific instruments and the detection of stable experimental phenomena often seem to survive radical theoretical and ontological changes.[35]

In recent times, the theory-ladenness vs. theory-freedom debate generated more nuanced standpoints. Even those who argue for the theory-ladenness of experimentation and scientific instrumentation, usually do not defend the circularity claims typical of some early anti-empiricist views and allow for the possibility to experimentally test our scientific theories (Chalmers 2003). A common strategy in both camps has been that of introducing different meanings of 'theory'. Heidelberger (2003) defends the possibility of theory-free experiments by introducing the category of 'theory-guidance'. This notion refers to how the theoretical background of the observer influences her disposition to make a particular observation. According to Heidelberger, theory-guidance does not determine in any way the *meaning* of observational sentences. Rather, it sits in between theory-ladenness and theory-freedom and admits of degrees and subcategories. In its minimal manifestation, e.g., in the case of exploratory experiments, theory-guidance borders with theory-freedom in Hacking's sense, because it does not influence experimentation with theoretical commitments, but only directs an improvement and expansion of causal knowledge of phenomenological regularities resulting from the instrument-phenomenon interactions. On the contrary, the more an instrument becomes embedded in a theory or paradigm, the more experimental observations become theory-laden. Yet, the general strategy of introducing different levels of theory has been criticised from philosophers arguing that all experimentation is theory-laden (Kroes 2003, Radder 2003), even if similar variants have been deployed even within this camp (e.g., Franklin 2005).

[34] Baird (2003: 47) convincingly argues for this conclusion based on his case study of Faraday's electromotor.
[35] Cf., for instance, Chang (2004: 51-52) on Feigl's view of experimental laws as robust mid-level regularities, and further references therein.

Overall, the historical development of the debate shows that early claims of theory-ladenness generalised from specific examples to argue for the pervasive influence of theoretical commitments in scientific inquiry. These claims usually come with an anti-realist flavour (i.e., scientific entities are nothing but theoretical constructs) and emphasise the threats of circularity, i.e., when experimental data are produced within a setting justified by the same theory that the data are supposed to corroborate. Claims of theory-free experimentation, particularly in Hacking's case, were made not only in the context of the '(re)discovery of the laboratory', but also as a reaction to both the dead-end of circularity and the anti-realist outcomes. In more recent times, the debate has shifted from issues of realism and focuses on providing examples that show the wide variety of theory-dependence relations and the complexity of interactions between theorising and experimenting, thanks to the dialectics generated by the introduction of different meanings and levels of theory.

If we look at the influence of this debate on Friedman's view of the relativised *a priori*, we notice that he still embraces a radical view of theory-ladenness, while sticking to a logical empiricist narrow view of experimentation as mere theory-testing. According to him, experimentation seems to be simply a way – plausibly the best one – to test theories by means of testing hypotheses, quite in line with the standard logical empiricist standpoint. Yet, experimentation is always theory-laden, because it necessarily requires the assumption of certain basic theoretical (i.e., conceptual) determinations – that is, the constitutive/relativised *a priori* principles – for experimentation to be meaningfully performed. If we look at the historical dimension, we see that this strong view of theory-ladenness also pervades Friedman's general perspective on scientific change, in that every radical change in the interaction with phenomena by means of instrumentation, experimentation, measurement, etc. is eventually justified by theoretical reasons and, ultimately, in terms of certain relativised *a priori* (theoretical) principles.

The highly idealised character of this radical view of theory-ladenness, although with no specific reference to the debate on coordination and constitutive elements, was already exposed by Galison's (1988) programmatic paper dealing with the issue of scientific change. Galison synthesises the history of the relationship between theory and experiment by means of three powerful philosophical images. The first two highlight the common search for universal dynamics of scientific development carried out by both the logical empiricists and their opponents. Both traditions share the common aim of attempting to identify a universal pattern of scientific evolution, where language plays a central role, in that it is a unifying element for the logical empiricists, but a central critical target for the anti-positivists:

> Kuhnian antipositivism and logical positivism […] share the search for a universal procedure of scientific advancement and a view that language and reference form the chief difficulty in the analysis of the experiment/theory relation. […] Both models have a well-established hierarchy that lends unity to the process of scientific work. True, they are flip-side versions of one another, but in their mirror reflections there is a good deal of similarity. (Galison 1988: 207).

In discontinuity with this search for a universal pattern, Galison rejects the assumption of a universally fixed hierarchical relation between experiment and theory. "Any model that we adopt, even a provisional one, ought to leave room not only for theoretical traditions, with breaks and continuities, but for experimental traditions with their own, equally complex set of internal dynamics" (Galison 1988: 208). This new image provides a mosaic framework that opens the possibility of a wide variety of relations between the categories of instrumentation, experimentation, and theories. A plurality of dynamical relationships holds between these components both diachronically – at different stages of scientific development – and synchronically – within different scientific enterprises – and different components might play the leading role of driving scientific change in different contexts. This composite picture allows for patches of continuity and discontinuity at the different levels of instrumentation, experimentation and theorising.

In my view, from Galison's new image a lesson must be drawn by all those who try to develop a view of coordination and constitutive elements, or of conceptual change in science more in general. This lesson is one of epistemic humility, in that any such view should not be guided by the search for universal patterns, but rather for similarities and common patterns. At the same time, Galison emphasises, once again, the importance of historical depth, in this case with respect to the history of philosophy, to the purpose of not getting biased by the strong influence that the category of 'theory-ladenness' had on recent traditions in philosophy of science, nor of forgetting the insight it gave us in reaction to its pitfalls. Therefore, I will approach my analysis by assuming that elements along different epistemic dimensions (theories, experimentation, instrumentation, etc.) may have a constitutive character in coordination.

1.3.3 Theorising, experimenting, and representing: Beyond a single unit of analysis

As we have seen, early perspectives on theory-ladenness seem to presuppose that the category of 'theory' and that of 'experimentation' are clearly distinct. However, as Morgan (2003: 232) highlights, it is very difficult "to cut cleanly, in any practical way, between the philosopher's categories of theory, experiment, and evidence". Still, a crucial difference between theorising and experimenting is usually identified in that the former aims at *representing* some empirical content by means of *symbolic tools*, whereas the latter aims at materially *manipulating* some *concrete physical interaction*. Philosophical characterisations of experimentation typically describe it as the material realisation of the interaction between some object of inquiry and some apparatus under controlled conditions in order to reveal a stable correlation.[36] In the debate on coordination and constitutive principles, the background stance, particularly in Friedman's perspective, is that theory and experimentation are clearly distinct, whereby theory – and, within it, certain relativised *a priori* principles – provides fundamental conceptual categories to set-up experiments and interpret their results.[37] In other words, experimentation can be a source of evidence inasmuch as it is performed according to criteria and representational tools prescribed and justified by theoretical means.

The fact that the distinction between theory and experiment is usually based on a clear divide between representation and manipulation is quite crucial here. What I challenge is the assumption that a representational character pertains only to symbolic representation, as contrasted with material manipulation, and is, therefore, a specific character of theories.[38] Following Gooding (2003), I consider experimenting as crucially involving representation, which is to be understood, however, as a selection and display of empirical content, rather than as the representation of this content by means of concepts or symbols and their relationships. Constructing any experimental set-up involves, at least, the abstraction (selection) of certain relevant empirical features and the idealisation (distortion) of some others in order to display a certain phenomenon (rather than another one) and, then, manipulate it, even in the case that the selection and distortion is operated by the instrumental apparatus and not by the epistemic subject.[39] It is in this sense that, in experimentation, "aspects of the world are *selectively redescribed* to make them amenable to manipulation according to rules that in many cases are now

[36] Cf. Harré (2003: 19), Morgan (2003: 216), Parker (2009: 487), and Radder (2003: 153).

[37] Clearly, this is peculiar neither to Friedman's view nor to other theory-focused views of constitutive elements, such as Stump's. On this specific issue, their views simply reflect the received standpoint, expressed earlier by the dominant syntactic view of scientific theories and, later, by the equally influential semantic view, on a clear-cut distinction between theory, which is abstract and mathematically expressed, and the messy realm of experimentation. Cf. Curiel (2019) for a sharp criticism of the semantic view on this aspect and for a convincing alternative pragmatic perspective. See also Ackermann (1985) for a consonant early view.

[38] In parallel, it could also be argued that a lot of theoretical activities are indeed based on manipulating symbols. However, this is not the place to develop that argument.

[39] Clearly, selection, distortion, display, and other tasks are often carried out via material manipulation but are not *justified* by manipulation itself. My argument moves from the rejection of a sharp divide between representing and intervening (cf. Hacking 1983). I believe that these two dimensions variously intersect. Representing is indeed a form of manipulation and, therefore, of intervention on something (even if, more typically, not on something *physical*), while intervention in experimental contexts minimally requires the (re)presentation of something, in the sense of selecting and displaying certain attributes of a phenomenon in an artificial context, even if selection, display, and manipulation are neither determined nor guided by theoretical representations.

implemented in machines" (Gooding 2003: 276).[40] However, this does not mean falling straight back to the thesis of theory-ladenness of experimentation *tout court*. Certainly, the appropriate level of abstraction, meaningful distortions, and accurate display are usually determined in the light of a theoretical background. Still, the extent of the representational character of experimentation and its dependence from a theoretical background is a matter of degree and, plausibly, it is influenced by the maturity of a discipline or research programme.

Once we understand that experimentation has a representational character of its own, and that this character depends to various degrees on the relationship between experimentation and theory, we have a guiding principle for constructing a conceptual space of theory-dependence in the form of a continuum (fig. 1.1).

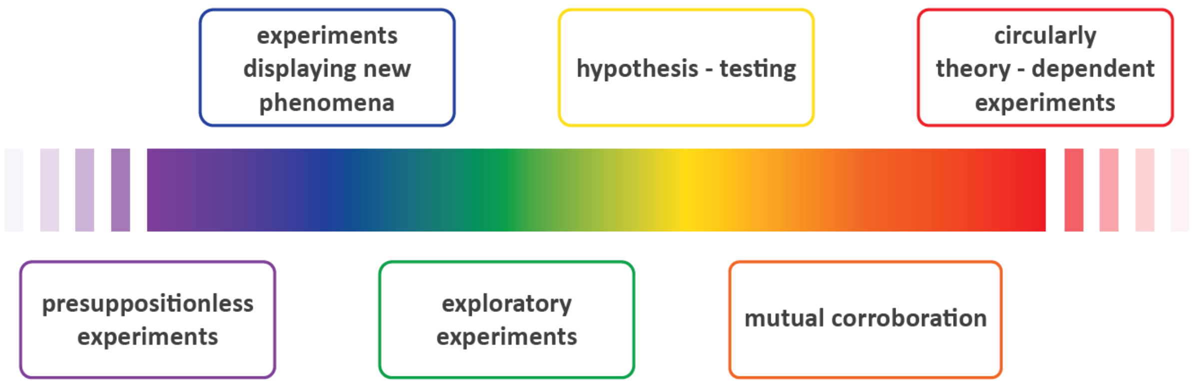


fig. 1.1: continuum of theory-dependence of the representational character of experimentation. Every item represents a different type of experiment according to the extent to which theory justifies the representational character of the experimental set-up

On the left-hand endpoint, I place presuppositionless experiments, to be conceived as ideal, rather than actual cases of theory-free experimentation. As the history of the debate on the theory-ladenness of experimentation shows, to any example of supposed theory-free experiments often corresponds a rebuttal, which highlights how, at a deeper level of analysis, these experiments do depend on some theoretical commitments. To use an example introduced in the previous section, Hacking's (1983) analysis of Newton's supposedly theory-free experimental observations of the dispersion of light is challenged by Radder (2003: 165-166), who shows at least three ways in which those observations were theory-dependent. However, this opposition can be dissolved by understanding how it conflates two dimensions: the representational character of experimentation, and its theory-dependence. As I claimed, experimentation, even if we simply consider the artificial set-up in which it is conducted,

[40] Emphasis mine.

requires the selection and display of certain features of a phenomenon under study, in contrast with its occurrence in nature. In this sense, it is difficult to imagine any experiment that does not involve the (re)presentation of anything. On the other hand, theoretical background often has a major role in justifying and fixing the experimental setup, and therefore in qualifying the representational character of experimentation. Yet, in some cases, the representational character of experimentation is almost entirely justified and fixed by the material apparatus (instrumentation and its qualitative characteristics) itself, and the performing of experimentation is justified by certain domain-independent assumptions, rather than by theory-specific ones.[41] In other words, in those cases, theoretical commitments have a minimal or no role in justifying the character of the experimental setup. Yet, experimentation is presuppositionless only to the ideal limit. This is because even if its representational character – i.e., the way the experimental set-up is constructed in order to select, display, and manipulate certain features of a phenomenon – is theory-independent, constraints to it can come from the material apparatus and from general assumptions required to make sense of the manipulation.[42]

By progressing towards the right endpoint of the continuum, the representational character of experimentation will be more and more justified by theoretical commitments. Here, I provide a non-exhaustive list of items reflecting the literature on the various kinds of experimentation and its roles, which can be placed on the continuum depending on the extent to which their representational character relies on theoretical assumptions, i.e., depending on their degree of theory-dependence.[43] These kinds of experimentation should not be conceived as mutually exclusive points on the continuum but rather as neighbourhoods with possible overlaps and interactions.

a. *Experiments displaying new phenomena*: some examples in the literature on theory-free experimentation belong to this kind. According to Baird (2003, 2004), Faraday's electromagnetic motor exhibited a rotary motion that could be created, reproduced, and manipulated despite the lack of theoretical understanding of the phenomenon. Still, the phenomenon was the product of a *selection* "of a suitable combination of electric and magnetic

[41] The literature on scientific experimentation usually mentions the domain-independent assumption of causal efficacy of a certain manipulation. In chapter 4 and 5, I will emphasise the role of further non theory-specific assumptions, such as approximation, in experimentation and along other epistemic dimensions.

[42] This consideration resonates Boon's (2004: 226) arguments against theory-free experiments *à la* Hacking "Hacking's argument that the existence of electrons has been proved since they can be used as tools is problematic since experimenters do not manipulate the electrons as such, but only an apparatus that is supposed "to spray electrons"".

[43] For example, Galison (1987) emphasises how some experiments are conducted to test the proper functioning of instrumentation. Placing this category of experiments within the continuum would require a long discussion of its own, which must be left to another project.

elements" (2003: 47), apparently performed in absence of theoretical commitments. However, this category does not include *all* experiments displaying, reproducing and manipulating new phenomena, but only those which do not require theoretical background.

b. *Exploratory experiments*: I adopt this category, even though it has been used with different meanings (cf. Steinle 1997, 1998, 2002, 2016, and Franklin 2005), exactly because it is quite general and leaves room for internal variability with respect to the degree of theory-dependence. In general, this category refers both to experiments conducted in the absence of a conceptual scheme and a local hypothesis guiding the research (thus, bordering with the previous category), and to experimentation that is not conducted to test specific theoretical hypotheses, yet is guided by some theoretical background. Therefore, the degree to which theoretical background justifies the representational character of the experimental setup is variable. In the case of domestication of phenomena, a natural phenomenon is simplified by giving it a new artificial set-up (Harré 2003). In theory-guided experiments, that is, experiments in which the theoretical background of the epistemic subject conditions their capacity to detect a certain phenomenon, theoretical background does not play a major role in justifying the construction of the experimental set-up, but it does guide the exploration of the physical interactions taking place within that experimental setting (Heidelberger 2003). Therefore, theory influences and directs scientific inquiry, but does not determine the representational character of the experimental set-up.

c. *Hypothesis-testing experiments*: this category refers to the general and rather uncontroversial case in which experimentation is conducted to test a specific and local theoretical hypothesis formulated against the backdrop of a larger theoretical background. To avoid circularity, it is crucial that the experimental set-up does not rely on theoretical constraints coming from the very theory that the experimentation is supposed to corroborate. Measurement is certainly not exclusive to this category, but it is typical of hypothesis-testing experiments to be aimed at detecting the presence of a certain property or to measure certain parameters.

d. *Experimentation resulting in mutual corroboration or coincidence*: According to Chalmers (2003), there are cases in which observational data and theory confirm each other at the same time, and he provides the example of how central arguments involved in the use of electron microscope were relying on novel theories of its workings and its interaction with the specimen under study. In these cases, the output of the experiment was interpreted in the light of the same theory underlying the working of the instrument in use. According to Chalmers (2003: 502), circularity was avoided by what he calls an 'argument from coincidence', that is, "that the match between theory, which here is a novel and hence previously untested theory about the interaction of specimen and instrument, and appropriate interpretations of the images on the micrographs serve both to confirm the theory and the interpretation of the data". In this case, only the mutual coherence between experimental output and theoretical justification of the experimental set-up, together with their coherence with further external assumptions, can dispel the threat of circularity.

We are now close to the right endpoint of the continuum which represents, again, an ideal case rather than a real one. I am referring to 'self-defeating' theory-ladenness, or, in my terminology, circular theory-dependence, that is, the case in which the representational character of the experiment is entirely justified by the very same theory which the experiment is supposed to test, while the theory to test does not receive support from other established experimental sources, either because it is novel or because there are no other developed experimental set-ups (thus ruling out coherence arguments). This ideal case might be understood, in regulative terms, as a negative limit by which circularity inevitably invalidates experimental results, although it might as well apply to real historical cases.

The analysis I have just provided shows that the relationships between the dimensions of theory and experimentation can vary a lot. When it comes to the issue of how scientists justify the coordination between certain abstract representational tools and the phenomena they are supposed to represent, my continuum shows that the view that only certain theoretical principles can provide justification for a form coordination may be too simplistic. Since experimentation has itself a representational character, but the extent to which theoretical background justifies this representational character varies, there may be further, non-theory-relative epistemic preconditions that justify the representational character of experimentation. In addition, the same may hold for further epistemic dimensions which variously interact with theory, such as measurement. Therefore, the assumptions that theories are the privileged unit of analysis and that local theoretical principles are the only form of constitutive elements in scientific inquiry may lead to overlook further fundamental epistemic components that perform a similar function along different epistemic dimensions.

In the rest of this dissertation, my goal will be that of providing a perspective on constitutive elements in science that is as neutral as possible with respect to the issue of the 'right' unit for the investigation of scientific inquiry. Therefore, my account of constitutive elements will not be tied to a specific unit, be it scientific theories, paradigms, or others. Yet, since any such view presupposes a distinction between the object of inquiry and the perspective from which it is investigated scientifically, I must commit, at least, to a broad notion of 'scientific framework'. By a scientific framework I refer to a network of people socially and economically organised in institutional structures and situated in a historical and geographical context, together with the common epistemic goals they pursue, the epistemic activities they perform and the material and symbolic tools by which they perform them, and the shared values that guide their inquiries. The precise borders of any such framework depend on the interests of one's philosophical, sociological, or historical inquiry. Depending on the purposes of a certain investigation *on* science, the historical span, the extension of the community of inquiry, the set of goals and activities, and the specific material and symbolic tools denoted by what counts as a scientific framework might vary. In addition, I do not deem any of the elements included in my definition as either necessary or sufficient to identify a scientific framework as a general construct, especially when it comes to the issue of understanding scientific change. On the lines of what Ankeny and Leonelli (2016) suggest with reference to their notion of 'repertoires', I take that any of the components included in the definition of a scientific framework, or any combination of them, can function as anchoring elements of the framework at different stages of development and, therefore, drive or prevent scientific change.

The constitutive elements at the centre of my analysis are those components that are cognitively required to implement the epistemic activities performed within a scientific framework, to the extent that they fundamentally contribute to justifying the coordination between abstract representations deployed by the framework and the phenomena they are supposed to represent. This does not exclude that social-institutional structure, historical factors, values and goals might themselves have a constitutive character, in that they can influence the acceptance or choice of the components I focus on. Yet, it is still worth making a distinction in epistemic function. Where the former can be understood as more directly constitutive, since they contribute to justifying a form of coordination, the latter are only indirectly contributing to it, by influencing the choice or acceptance of certain (first-level) constitutive components. This dissertation only focuses on the first-level components.

CHAPTER 2

2.1 Coordination, presuppositions, and constitutive elements: a working definition

So far, I have introduced the main contemporary views of the constitutive *a priori* in science, their similarities and differences, and a few general methodological assumptions of my own approach. In this chapter, I will present a novel degree and context-sensitive account of the constitutive character of certain epistemic components in scientific practice. In this section, I will set out to focus on the notions of 'coordination' and 'presupposition', which I deem crucial to provide a general working definition of constitutive elements. Then, in section 2.2, I will provide a detailed analysis of Friedman's account and of two of his examples, I will spell out some limitations of his view discussed in the literature, and I will outline my own criticism. In section 2.3, I will discuss some background assumptions of my own account and I will introduce its core features. In section 2.4, I will reframe one of Friedman's examples in terms of my own account, to show its advantages compared to Friedman's, Quine's, and Wimsatt's views.

As I mentioned in chapter 1, discussions on the role and character of the constitutive *a priori* in scientific inquiry have often been related to the issue of coordination. Scientists use abstract tools, such as mathematical symbols, concepts, propositions, graphs, etc. to represent the empirical world, that is, to describe, explain, understand, and theorise over widely different domains of concrete phenomena. For these abstract tools to represent the domain of interest and, thus, lead to successful theories and interventions, the referential relation between these abstract representations and the phenomena they are supposed to represent must be established and justified, since it is not given in the first place. In other words, some form of coordination between phenomena and their abstract representation must be in place for these representations to be meaningful and, therefore, to enable the pursuit of some epistemic goal.

In section 1.1.2, I pointed out that philosophers in the empiricist tradition offered different answers to the issue of coordination. Contemporary debates on coordination build on these classic views already by setting up the problem of coordination with a specific focus, and not only in the light of the solutions that they offer. For instance, Friedman (2001) sets out to address coordination in its general form, as the problem of connecting the abstract level of mathematical formulations, typically used to express scientific laws, with the concrete one of real phenomena. On the other hand, Chang (2004) and van Fraassen (2008) frame the problem of coordination as an issue of circularity between the meaning of quantity concepts and the procedures measuring those quantities, thus narrowing down the scope of the problem itself. As far as the respective solutions to the problem are concerned, Friedman resorts to constitutive *a priori* principles to solve the general version of the problem he tackles (more in the next section), whereas contemporary epistemologists of measurement consider empirically-guided practice and pragmatic considerations as the main sources of coordination between quantity terms and measurement procedures.

In my view, the issue of coordination can present itself in different forms, some more general, others more local; some concerning epistemic items such as scientific laws, some others concerning theoretical terms, yet others concerning empirical data (e.g., measurement outcomes), etc. Their common core, however, is that coordination is to be understood as a *process*. Through this process, the epistemic subjects working in a scientific framework establish and provide the justification for the referential relationship between an abstract representational tool and a represented item, where the extent to which a representation is abstract and the represented is concrete – or closer to the empirical world – is a matter of degree and context. Once a form of coordination is established and the state of knowledge in that domain is, so to speak, settled, coordination becomes part and parcel of the fixed referential relationship, which can, nonetheless, always be subject to modification or replacement in the light of scientific developments. Given this understanding of coordination, my goal is to identify those epistemic components that have a constitutive character, i.e., that contribute to justifying a form of coordination, in that they provide certain conditions for the identification of the referential relationship between objects of scientific inquiry and their abstract representation.

These components might be qualitatively heterogeneous and, in most cases, may be identified only by delving into the nitty-gritty details of scientific practice. One thing seems quite straightforward: the notion of 'constitutive', as applied to specific conditions for empirical inquiry leading to a successful coordination, is closely connected to that of 'presupposition'. As we saw, some, even rough, form of coordination must be in place, so that a certain representation can be deployed to pursue an epistemic goal. In other words, some form of coordination between an abstract representation and its concrete referent is presupposed, before the epistemic subjects can engage in some scientific activity. Whatever more precise connotation should something possess to count as constitutive within a scientific framework, it seems uncontroversial to say that it will have to be presupposed, i.e., held *as* true or valid for the formulation, applicability, and (sometimes) testability of the tools for scientific representation and intervention required by that framework.

Attempts at characterising the notion of presupposition were made in the logical empiricist tradition, to pinpoint its role in the logical structure of theories.[44] According to van Fraassen (1968), A presupposes B if and only if $A \rightarrow B$ or $\neg A \rightarrow B$. On the other hand, if $\neg B$ is the case, A does not have a truth-value. As van Fraassen points out, presupposition is a trivial semantic relation, if we endorse the principle of bivalence, i.e., that every sentence is, in any possible situation, either true or false. Brittan (1978) deployed van Fraassen's semantic account to interpret the Kantian notion of constitutive *a priori*. According to Brittan, something counts as a constitutive presupposition if it determines the possibility of an object to be experienced, by somehow positing the possible objects of experience. This means that a presupposition is constitutive inasmuch as it *posits* the entities the absence of which would (quite banally) preclude further (empirical) investigations and specifications of those very entities themselves. In this sense, the (empirical) statements that are established by means of those further investigations and specifications assume *that* presupposition as a given.

However, Brittan's interpretation would regard as 'constitutive' any sentence that posits the existence of (scientific) entities, thus leading to a very weak understanding of the constitutive character. Yet, this permissive view can, with some further qualification, provide a good starting point for characterising the notion of 'constitutive' suitable to identify the components involved in the justification of a form of coordination. Here, I provide a working definition of 'constitutive elements' that will receive further specification in the course of this chapter, but which expresses the general common characteristics shared by these heterogeneous kinds of epistemic components:

> **Constitutive elements:** *Given the scientific framework of inquiry under scrutiny, some of its elements are constitutive in that they are preconditions enabling the epistemic subject(s) to perform the epistemic activities required to operate within that framework, insofar as they contribute to justifying the referential relationship between certain representational tools and the phenomena they represent*

In other words, these elements play a crucial role in establishing the coordination between the representational resources needed to perform successfully the epistemic activities of a scientific framework, and the phenomena that those resources are supposed to represent. Rather than being features merely inferred from empirical input, these elements are better characterised as enabling conditions for our scientific representation of – and engagement with – the empirical world. This characterisation might include instances of several types of epistemic components, at different levels

[44] For some classic papers on the notion of presupposition, see Strawson (1950), and Hintikka (1959).

of generality (e.g., theoretical principles, measurement procedures, scientific instruments, etc.). In fact, these elements may perform distinct specific functions, and even be hierarchically organised.

This working definition seems straightforwardly applicable to linguistic elements, particularly to statements belonging to scientific theories. This would make my definition continuous with notions such as 'coordinative definitions' and 'constitutive principles' as usually discussed in the tradition of logical empiricism and neo-Kantian perspectives on science. However, my definition substitutes the linguistically-oriented notion of 'presupposition' with the more open 'precondition', thus accommodating the possibility to analyse the constitutive character of non-linguistic epistemic components.[45] For instance, a material component of an experimental setting can be a precondition for achieving coordination and, thus, justification for some knowledge claim, but cannot be said to be a presupposition in a propositional sense. In addition, in light of my definition, some components involved in coordination that are hardly characterizable in terms of the traditional interpretations of the *a priori* and would rather be considered as empirical by more naturalistic approaches, will count as constitutive to some degree. The degree and context-sensitive account that I develop in section 2.3 provides the means to account exactly for this complexity. Consequently, I clarify how certain features of the classic interpretations of the *a priori* are loosened in my approach, compared to some contemporary perspectives. More specifically, I focus on Friedman's view, which is going to be my main critical target for the rest of the chapter.

2.2 The limits of Friedman's account

2.2.1 External critique of Friedman's view

In section 1.2 of the previous chapter, I briefly reviewed the contemporary perspectives on the constitutive *a priori* in science. In my review, I started off with Friedman's account, since I consider it the default position in the contemporary debate for two main reasons. Firstly, Friedman's work is the most referred to in the literature on constitutive principles in science. Secondly, it develops a *general* notion of constitutive principles that impacts on various philosophical and scientific issues, such as the role of philosophy with respect to the sciences, or the issue of conceptual change across scientific revolutions.[46]

[45] Cf. below, chapter 4. Yet, given the traditional focus of discussions on the constitutive *a priori* in philosophy of science on statements of scientific theories, I will start my discussion from cases of theoretical principles.
[46] However, several works on the topic of constitution with respect to specific issues, particularly in modern physics, do not take Friedman's view as a point of reference (cf. Bitbol et al. 2009, Castellani 1998).

Friedman's encompassing perspective has influenced both historical and philosophical research on the role of *a priori* principles in science, particularly – but not only – in space-time theories (e.g., DiSalle 2006, Ryckman 2005, Stöltzner 2009). Yet, many points have been raised against Friedman's picture of scientific change as characterised in *Dynamics of Reason* and in following works (Friedman 2010, 2012). To start with, although his account of conceptual change in science assumes Kuhn's contrast between normal and revolutionary science in the background, it portrays a rather orderly change despite scientific revolutions. In fact, the passage from one scientific framework or paradigm to the following – in Friedman's language, from one set of relativised *a priori* principles to the following – happens through a series of natural transformations such that a new framework supersedes the previous one, which maintains its validity as limiting case.[47] However, this idealised picture does not seem to give justice to the actual state of science, often characterised by multiple competing scientific frameworks working in the same domain, episodes of messy transformations, and scientists capable of shifting from one framework to the other for pragmatic purposes.[48] Secondly, Friedman's analysis seems to assume that the sort of constitutive elements he identifies in space-time theories should be taken as a blueprint for the understanding of the structure of scientific knowledge in other sciences, too (Friedman 2001: 117-129). In addition to this, he claims that the principles that are constitutive of space-time theories represent fundamental assumptions for all empirical inquiries.[49] However, very few scientific frameworks can satisfy the demands of this strongly prescriptive account, based on a single set of universally-valid constitutive principles at any given stage of scientific inquiry.[50] All these

[47] This point is important in the light of Friedman's understanding of his own project as a historicised version of the Kantian *a priori*. The supposedly natural character of the transformations lies in that they are a necessary consequence of the growth of science. This aspect has been criticised, for instance, by Buzzoni (2013a, footnote 1): "A relativized *a priori* whose changing depends upon the growth of experience and knowledge is by no means Kantian, because it cannot answer, in principle, any *quaestio juris*, i.e., it can neither justify the validity of an inference nor the truth of a knowledge claim".

[48] Cf. Chang (2008). This is certainly the case for the life sciences, where synchronic epistemic pluralism is ubiquitous (cf., e.g., Leonelli 2009, Morrison 2011), but also in electrochemistry and atomic chemistry (Chang 2012). Even in the case of space-time theories, it is often noted that Newtonian mechanics is commonly taught and used for a wide variety of purposes, and that scientists often work across frameworks. A commonly cited example is that of geographic information systems (GIS), which combine elements from Newtonian mechanics, relativistic physics, and quantum physics (cf. Chang (2012) and Winther (2015)).

[49] See above, section 1.2.2. Stump (2015) criticises Friedman on this point, based on Norton's (2010, 2014) account of induction. Rather than very general assumptions on nature in general, scientific inquiries are usually grounded by specific presuppositions (cf. also Padovani, 2017). In the words of Love (2013: 335): "Changes in concepts due to empirical advances are localized and therefore tend to exhibit distinctly regional affects in scientific theorizing, and empirical testing, because tests are localized to a particular MIS [material inferential structure] such that positive or negative results do not echo into every other theory structure".

[50] Cf. Chang (2008), Mormann (2012), and Suàrez (2012) on this point. Further critical stances on different issues include Everett (2015), Ferrari (2012), Mormann (2012), Padovani (2011, 2015), Shaffer (2011) and Uebel (2012).

issues can be viewed as external critiques of Friedman's perspective, which expose the limits of Friedman's account with respect to goals other than the ones he set himself. In my view, these criticisms can be better understood by taking a closer look at some specific features of Friedman's interpretation of the constitutive *a priori*, and at its own internal limits. That requires focusing on the two least controversial examples of constitutive *a priori* principles that he develops: the laws of motion in Newtonian mechanics and the light principle in relativistic physics.[51]

2.2.2 Friedman's examples of the laws of motion and of the light principle

Friedman's analysis of the structure of Newtonian mechanics is inspired by Kant's theory of pure natural science, which took mathematical physics and astronomy as paradigmatic empirical sciences and Newtonian mechanics as the central example of his inquiry on how natural science is possible.[52] According to Friedman, the revolution brought about by Newton's scientific advancements consisted of three different innovations: "a new form of mathematics, the calculus, for dealing with infinite limiting processes and instantaneous rates of change; new conceptions of force and quantity of matter embodied and encapsulated in his three laws of motion; and a new universal law of nature, the law of universal gravitation" (Friedman 2001: 35). Certainly, Newton developed these innovations to tackle the same epistemic issue of providing a unified theory of motion for both terrestrial and celestial phenomena. Yet, Friedman highlights that these three elements fulfil different epistemic functions, therefore giving rise to asymmetries internal to the structure of Newtonian physics. For example, in the second law of motion the notion of 'acceleration' is conceived as instantaneous rate of velocity, i.e., the instantaneous rate of change of position. Without a mathematical tool like the calculus, tailored to deal with instantaneous rates of change, the second law of motion could not even be formulated. Friedman makes the same point also with respect to the relationship between the three laws of motion and the law of universal gravitation. To make empirical sense of gravitational physics, we require the concept of inertial frame, that is, of a frame of reference relative to which the accelerations of two bodies as predicted by the law of gravitation occur. In such a framework the laws of motion hold, as in the (approximate) case of the centre of mass frame of the solar system. In other words, without the concept of inertial frame and, therefore, without the three laws of motion, the law of universal gravitation could not be formulated nor empirically tested: "we would simply have no idea what the

[51] Friedman's reconstruction of the passage from Newtonian framework to the framework of general relativity is contested. Howard (2010) and Pitts (2018) convincingly argue against Friedman's characterisation of the equivalence principle as constitutively *a priori*, as interpreted by Friedman, within the general theory of relativity. See also Everett (2015). These criticisms, however, do not undermine the fact that the light principle can be adequately characterised as constitutively *a priori* in Friedman's sense (cf. also Friedman 2009).
[52] Cf. Friedman (1992).

relevant frame of reference might be in relation to which such accelerations are defined" (Friedman 2001: 36).

This example clearly illustrates Friedman's understanding of physical theoretical frameworks as comprising three levels, a mathematical, a mechanical, and an empirical one, where the mechanical part establishes a coordination between the mathematical and the empirical parts. In Newtonian physics, the three laws of motion are the mechanical part, in that they establish the connection between the mathematical part, supplied by the mathematics of the infinitesimal calculus, and the empirical (or first-order) part, which is the domain of phenomena described by the law of universal gravitation. Within the context of a theory, the mathematical part and the mechanical part are together *constitutive* of the empirical part. The mathematical and the mechanical parts are therefore *a priori*, but only in the sense that they are constitutive of the fundamental theoretical concepts of a theory, since they provide basic elements of the conceptual machinery that is required to formulate and test the first-order empirical law of universal gravitation. In this way, Friedman manages to offer a more differentiated and powerful epistemological view than Quine's holism, since a distinction in function can be traced between the internal components of a theory, independently from their degree of entrenchment. For example, the calculus was not a more entrenched or universally accepted scientific tool than the laws of motion at the time in which Newton used in his scientific work. At the same time, Friedman's account does not fall victim of the same issue that Kant faced, i.e., that the *a priori* components cannot change across time. In Friedman's view, here inspired by the early Reichenbach's (1920/1965) perspective, the constitutive components change during profound conceptual revolutions, which bring about a novel scientific framework. For instance, the transition from Newtonian mechanics to relativistic physics is marked by a major conceptual shift, which is determined by a change in the framework's constitutive principles.

The conceptual shift characterising the transition to the relativistic framework is essentially a change in the "network of inferential evidential relationships" that must be presupposed in order for an observation or measurement to count as evidence for a theoretical claim (Friedman 2001: 85). In other words, the constitutive components of the new framework, i.e., the logical-mathematical principles and the coordinating (mechanical) principles, provide themselves the justification for the inferential "jump" from a certain piece of evidence to a theoretical claim within a certain framework. According to Friedman, it is in virtue of the constitutive role of the Riemannian theory of manifolds (mathematical part) and of Einstein's coordinating principles (mechanical part) that the advance of the perihelion of Mercury can count as an evidential reason to accept Einstein's field equations (empirical part), rather than it being just a pragmatic or instrumental reason. This piece of evidence, in contrast, cannot itself justify the acceptance of the coordinating principles of the framework, such as the light principle. However, this does not mean that the light principle cannot be empirically tested at all, as it was shown already in the 1880s by Michelson and Morley, who devised an experimental apparatus to test the invariance of the velocity of light in different inertial frames. Still, their experiment was only demonstrating that the behaviour of light showed no detectable changes due to the motion with respect to the aether. The experiment did not itself demonstrate the invariance of light in *all* inertial frameworks and, therefore, could not be an empirical proof justifying the epistemic function given to the light principle in the relativistic framework: "the Michelson-Morley experiment can in no way be viewed as an empirical test or "crucial experiment" of special relativity *with respect to its theoretical alternatives*; and it is not a test, for precisely this reason, of the relativistic light principle" (Friedman 2001: 87).[53] In fact, what Einstein did was to 'elevate' the empirically discovered fact – that inertial motion did not produce detectable changes in the behaviour of light – from the status of inductive generalisation to the status of conceptual presupposition, by which he was then able to develop a new framework for space, time, and motion alternative to the classical one.[54] In other words, he gave the light principle a novel epistemic function which makes this principle constitutive within the relativistic framework.

2.2.3 Towards a new approach: spelling out Friedman's gradualism

The analysis of these examples illustrates that Friedman embraces an interpretation of the *a priori* according to which those elements that perform a coordinating function count as *a priori* even if they have some empirical content. In fact, neither the Newtonian laws of motion nor the light principle are *a priori* in the traditional sense, according to Friedman, since they evidently say something about the world. Still, these principles contribute to making up the structure of the first-level empirical theories under scrutiny, respectively, the law of universal gravitation and the special theory of relativity, without being merely the product of inductive generalisations. These principles are constitutive in that they provide fundamental concepts and the conditions of testability of other (networks of) empirical statements, while they cannot be themselves empirically tested from within their own framework. Therefore, these principles constitute, that is, conceptually determine, the first-order empirical laws, where 'first-order' refers to those lawful empirical generalisations that directly structure the concrete empirical phenomena.

[53] Emphasis mine.
[54] Cf. Friedman (2009).

As we saw, Friedman's interpretation of the constitutive *a priori*, by which it can (but does not have to) have empirical content, is not the only one available.[55] As emphasised by the label 'first-order' applied to empirical laws, Friedman's interpretation of the *a priori* is crucially supported by his view of the structure of physical theories in three levels (mathematical, mechanical, empirical). This view allows him to accommodate his interpretation of constitutive *a priori* elements as working between the level of mathematical structures and that of concrete phenomena. At the same time, it solves the problem of coordination, i.e., of how to coordinate abstract mathematical structures empty of content with empirical phenomena, so that the former can adequately represent the latter. The coordination is established by the constitutive elements at the mechanical level.

As pointed out by Tsou (2010), Friedman's distinction of three levels internal to physical theories seems to allow for a distinction between different degrees of apriority or *constitutivity*. The mathematical principles are, in fact, *a priori tout court*, in that they cannot be directly tested or disconfirmed in the light of experience. The mechanical (coordinating) principles are, in a sense, *less a priori* than the mathematical ones, in that they partake of empirical content. Following the same direction, the infinitesimal calculus seems more constitutive than the laws of motion, since it provides the concepts to express the laws of motions themselves, while both the calculus and the laws of motion are required to formulate and test the law of universal gravitation.

This gradualism, however, is not spelled out explicitly by Friedman. Evidently, the gradualism resulting from tying this interpretation of the constitutive *a priori* to a specific analysis of the structure of physical theories would prove very hard to defend in other scientific contexts. It seems plausible that Friedman's interpretation within the context of space-time theories implies a kind of gradualism that may be peculiar only to highly mathematised physical sciences, or even just to a specific subset of physical theories adequately described by the three-level structure that Friedman presupposes. In fact, the gradualism intrinsic to his analysis of the structure of space-time theories is implicitly mirrored by the degrees of constitutivity of the principles located at different levels of the theoretical structure. However, once it is spelled out, this characterisation seems in tension with the claim of universality and fundamentality of a supposed single set of constitutive *a priori* principles that Friedman identifies for each stage of scientific development. In addition, as I already highlighted, it has been argued that there are components of scientific inquiry that have a role in coordination, yet not along the dimension of theorising, as in the case of measurement procedures (van Fraassen 2008). Consequently, even the issue of coordinating our mathematics with empirical phenomena might not be solved only by means

[55] Cf. above, chapter 1, sections 1.2.1, 1.2.3.2, and 1.2.3.3.

of the theoretical principles that Friedman allows for (Padovani 2015, 2017), even if we embrace his three-level analysis.

These issues internal to Friedman's account, together with the general bias of Friedman's interpretation towards mathematical physics, provide substantial motivations to consider Friedman's view of coordination and constitutivity only adequate, at best, to represent the two examples described in the previous section. Yet, it is possible to elaborate an account of constitutive elements that, on the one hand, acknowledges and pushes the boundaries of the gradualism implicit in Friedman's view and that, on the other, allows for accommodating constitutive components different from those Friedman focuses on, by decoupling the gradualism from his analysis of the structure of physical theories.

2.3 Building a new account of constitutivity

2.3.1 Units of analysis, focus on practice, and hints from the evolutionary perspective

As I pointed out, Friedman's view of coordination and his interpretation of constitutive principles are developed against the backdrop of Kuhn's view of paradigmatic shifts in the history of science. According to Friedman, Kuhnian paradigms are constituted by relativised *a priori* principles that provide the conceptual framework within which the problem-solving activities of normal science can take place. These constitutive principles change during profound conceptual revolutions, which bring about a novel paradigm.

A point worth noting is that the replacement of crucial conceptual tools for scientific theorising does not happen just at the macro-level of revolutionary paradigm-change. In fact, what can rightfully count as a paradigm might largely vary from those of space-time physics which Friedman takes into consideration.[56] Some conceptual or theoretical changes might appear as revolutionary in a restricted domain, but from the point of view of a larger domain they could just be considered as rather minor modifications. As highlighted by Wimsatt (1987: 17), "we tend to dichotomize a continuum of cases into two alternative cases – 'revolutionary' and 'normal' science, but the continuum is there nonetheless." A crucial assumption of the account that I will develop in this section is the acknowledgment that some epistemic components have a constitutive character even if they do not arise from major paradigmatic shifts. Kuhnian paradigms are not necessarily the only relevant meta-unit of analysis for the investigation of the constitutive components in the achievement of a form of coordination. Thus, my account intends to overcome the bias towards the extreme of radical

[56] Kuhn himself (1970: 177 and 180-1) allowed for a great deal of variability with respect to the scope and degree of generality of a scientific framework to count as a paradigm.

conceptual replacements, and to make sense also of piecemeal ones, in so far as the latter contribute to the coordination of abstract conceptual structures and concrete phenomena along any epistemic dimension (theorising, experimenting, measuring, etc.).

This bias can be controlled by considering alternative meta-units of analysis for the study of scientific change, and by taking cues from analyses of conceptual change in science that ran in parallel to the tradition of integrated history and philosophy of science. One example is that of evolutionary accounts of science, that characterise scientific inquiry as an on-going selection process (Campbell 1974, Hull 1988, Popper 1972). Clearly, the naturalisation of the process of conceptual change in science is a core thesis of most evolutionary accounts that an HPS tradition would hardly embrace.[57] Still, exchange and contamination between the two traditions could be fruitful (and it is still greatly underexplored), considering that evolutionary accounts and the HPS tradition share the perspective of *provisional and situated knowledge*: "knowledge acquisition [is to be considered] as a temporal process […] carried on by fallible human beings whose careers have an inevitable temporal limit" (Hull 1988: 7).

The latter point leads to a second crucial assumption of my new account. As I argued in section 1.3.2, it is necessary to address the issue of coordination and constitutive elements by focusing on the entire process of scientific inquiry, and not just on scientific knowledge as a final product. The complexity brought in by the historical dimension should go hand in hand with the attention to the actual practices of scientists with respect to the forms of coordination that they establish or assume. The dimension of practice should certainly encompass theorising as a complex scientific activity, but it should be viewed as complementary, rather than alternative, to considering theories and principles in the abstract, as something settled and now belonging only to the context of justification.

2.3.2 Borrowing concepts from Wimsatt: Generativity and quasi-independence

The focus on piecemeal conceptual replacements characterising many cases of conceptual change in the practice of scientists, together with an evolutionary approach broadly understood, are two core features of the science-as-engineering view developed by Wimsatt (2007). I now briefly introduce the general traits of his perspective, since I will use two notions from his account, 'generativity' and 'quasi-independence', to construct my own account.

Wimsatt (1987, 1999) elaborated a naturalistic view of scientific change that, at the same time, could also replace the analytic/synthetic divide. His view sets out from an analogy between evolving systems

[57] Cf. Callebaut (1993, ch. 7) for an in-depth discussion.

in the realm of biology and scientific inquiry as characterised by an organic development. Within this development, certain *generative systems* have a distinctive function in that they provide developmental constraints along the path of scientific progress. These systems are characterised by a high degree of *generative entrenchment*, which is proportional “to the number of downstream features which depend upon them" (Wimsatt 1987: 5). The notion of generative entrenchment originally comes from Wimsatt's treatment of biological systems and is a measure of how much of the generated structure or activity of a complex system depends upon the presence or activity of that entity (Schank & Wimsatt 1986). When applied to the analysis of scientific knowledge, the notion of ‘generative entrenchment’ is not identical to that of ‘entrenchment’ presupposed by Quine’s web-of-belief. According to Wimsatt, generatively entrenched features are resistant to replacement in virtue of the conservatism typical of early features in developmental paths. In other words, features that appear early in the developmental path of scientific inquiry and are not replaced at an early stage, become generative of many other downstream features, which develop out of them or against the backdrop of the constraints that these early features put into place by making certain developmental paths unviable. Consequently, the replacement of highly generatively entrenched components is very costly, given that it has disruptive consequences on a large scale, and happens very rarely. This view implies a cumulative picture of science, and the analogy with the conservatism of early features in phylogeny clearly assigns a central role to the historical character of generative entrenchment. On the other hand, Quine is often unclear whether the correct understanding of his own notion of ‘entrenchment’ has anything to deal with the historical development of science or must be understood simply on a logical level.

In addition to generativity, according to Wimsatt (1987: 8), the organization and interactions between scientific theories and disciplines are characterised by *quasi-independence*, which is essential for evolution to proceed at a reasonable rate. This means that each scientific ‘niche’ preserves a significant degree of autonomy from the development, also conceptual, of other disciplines, fields, and theories. Without this independence, scientific theories, sub-disciplines, etc., would have to bear too many constraints to evolve as fast as they do. Generativity is, indeed, a property that admits of degrees. Therefore, the replacement of components less generatively entrenched than others will lead to less disruptive consequences on the relevant system. However, what counts as the relevant system affected by the replacement does not, according to Wimsatt, boil down to a Quinean web-of-belief. The notion of ‘quasi-independence’ captures the fact that the connections between bodies of scientific knowledge are highly variable, can be found at different levels of generality, and do not preserve what Quine would call ‘the total system of knowledge’ from possible (and actual) inconsistencies, thus leading to cases of inconsistent epistemic pluralism even within the same domain. In my view, these inter-theoretical and inter-disciplinary connections, rather than by a homogeneous web where inner components have a higher degree of entrenchment, are better described by the image of a variable landscape: they are often hierarchical and non-homogeneous in places. This picture certainly does not deny that our bodies of scientific knowledge often show a significant level of interconnectedness and integration. However, it emphasises the fact that some areas are patchy and discontinuous, and – more importantly - that the outcome of science often consists of accounts that are non-reducible to, and are overlapping with, one another: "The structure of scientific knowledge *is* reticular, but [...] the web is, and should be, *multi-layered* in places" (Chang 2012: 259).[58]

Still, Wimsatt's evolutionary-developmental account shares with Quine the emphasis on the *pragmatic cost* of revision as justification for the choice of replacing certain components rather than others. According to Wimsatt, in the day-by-day workings of science, the prevailing tendency is to replace components at the smaller level possible rather than to shake foundational elements. In his view, some elements certainly have a peculiar epistemic status, given their high degree of generative entrenchment, but they do not necessarily coincide with those encapsulated by Kuhnian paradigms, nor is their replacement a prerogative of revolutionary science. Yet, Wimsatt's naturalistic account has some limitations that he himself acknowledges. For instance, the general point that deeply generatively entrenched features bear a great amount of dependence relations accumulated along the developmental path, is not accurately mirrored by many episodes of conceptual change in the history of science. Abstraction, generalisation, and reinterpretation can very easily make parts of a theoretical structure more generatively entrenched, without the dependence relations being gradually accumulated in time. What is more, cases of profound and relatively quick conceptual modifications can be hardly made sense of within the evolutionary-developmental analogy.[59] Yet, Wimsatt's view provides some useful conceptual resources to make sense of the replacement of epistemic components at different scales. The notion of 'generativity', albeit decoupled from the feature of entrenchment, and that of 'quasi-independence' will be central to my new account.

58 Italics mine.

59 Further disanalogies concern the low cost of producing conceptual alternatives as opposed to biological alternatives, and the possibility and different features that characterise the adoption of such conceptual alternatives (cf. Wimsatt 1987: 26-27).

2.3.3 Core features: Quasi-axiomaticity, generative potential, empirical shielding

In this section, I introduce the core features of my novel account of constitutivity. This account is degree-sensitive because it allows for different epistemic components to be constitutive to various degrees. It has a quantitative character, since constitutivity is expressed in terms of certain degree-sensitive parameters, but the values of these parameters are assessed on the basis of contextual analyses of the scientific framework under investigation. *From within* that scientific framework, certain epistemic components will show a certain degree of constitutivity compared to others, given their role in establishing a form of coordination. Therefore, my attempt at quantification is not to be taken as a strong formalisation striving to be universal, but as an attempt to abstract away from the individuality of specific case studies and to set a common quantitative (in the ordinal sense) ground for comparison. In the next section, I translate Friedman's case of Newtonian mechanics in terms of my own account. I will then test my account with new examples starting from the next chapter, to show how the gradualism can help making sense of the constitutivity of certain epistemic components beyond the limits set by Friedman's interpretation. In addition to the quantitative framework, I will introduce qualitative distinctions between different types of constitutive components starting from chapter 4.

Recall my working definition of constitutive elements:

> **Constitutive elements:** *given the scientific framework of inquiry under scrutiny, some of its elements are constitutive in that they are preconditions enabling the epistemic subject(s) to perform the epistemic activities required to operate within that framework, insofar as they contribute to justifying the referential relationship between certain representational tools and the phenomena they represent*

As I already mentioned, this characterisation is rather broad and might include instances of several kinds of epistemic components, across different epistemic dimensions. A powerful account should be able to accommodate the fact that there are specific differences between them, but it is by adding contextualised and historically situated analyses of specific cases that these differences should be accounted for.

The three features that express the core of the notion of constitutivity and enable its quantitative representation are the following:

Quasi-axiomaticity. Certain epistemic components have a feature often attributed to conventions or analytic truths, in that they introduce standards and constraints for describing and investigating one or more domain(s) by means of specific concepts or procedures.[60] They are held constant, while allowing the other variables that comprise the framework to be tested or changed. Their degree of quasi-axiomaticity depends on the extent to which they are fixed elements within a framework of inquiry, the prefix 'quasi' denoting the absence of a strict axiomaticity requirement, i.e., the requirement that these elements should be rigorously and explicitly defined or standardised.[61] In my account, quasi-axiomaticity refers specifically to the fixity of the *epistemic function* of a component within a framework, rather than the precise form or semantic content of that component. Clearly, form and semantic content can exhibit variability both across sub-frameworks – as in the case of different research groups within a research program using, for instance, a slightly different mathematical formulation of a principle – and across time, as in the case of variations in the formulations of a concept at different stages of development of a scientific framework. Depending on the scope of one's epistemological analysis, considerations of variability will determine, in this sense, what counts as the *same* (constitutive) component. Once the component has been identified, its degree of quasi-axiomaticity expresses the extent to which its epistemic function within a framework is held fixed, so that other parts can be empirically tested, modified, or replaced.

Generative potential. Generative potential identifies the amount of dependence-relations that one or more epistemic components can support with respect to other components of the scientific framework under analysis. Since constitutive elements enable the emergence of fundamental *conceptual* resources for investigating and representing an empirical domain, more domain-specific components (e.g., empirical laws, empirical models, models of data, explanations, complex observational protocols etc.) will themselves partially depend on these elements. This means that the most paradigmatic cases of constitutive elements will present a high degree of generative potential, in that they support a large amount of dependencies that ground other conceptual resources employed in a framework of inquiry and allow for the emergence of a large enough conceptual space to generate further empirical developments (hence the term 'potential').[62] This parameter clearly shows similarities with Wimsatt's

[60] Note that these components can themselves be concepts, but do not have to, as in the case of material constraints, which I analyse in chapter 4. What is important is that they constrain the use of certain conceptual resources.

[61] Despite the lack of the requirement of strict axiomaticity, one of the primary tools that allows for fixing some element is certainly formalisation. Formalised languages (especially, but not exclusively, mathematics) and procedures, which extend beyond the mere use of symbols to, e.g., graphic representation, or standardisation of measurement procedures and tools, enable to set these elements fixed, often by incorporating a certain amount of idealisation.

[62] In the case of constitutive elements that are propositional in character, these dependence-relations do not hold in virtue of the *truth* of these elements since this would already presuppose a one-to-one correspondence between such conceptual elements and the respective (classes of) objects of empirical inquiry.

notion of 'generative entrenchment'. However, it decouples the feature of generativity from that of entrenchment which, in Wimsatt, necessarily results from a historical-developmental process, and ties it to the potential of a component to ground epistemically a certain amount of dependence-relations, in abstraction from the actual developmental process of a scientific framework.[63] The result is a notion that is close to the constitutive feature in Friedman's sense, but without the component of apriority and universality that are embedded in Friedman's interpretation. Consequently, generative potential might seem to be the most fundamental among the three features identifying constitutivity. However, this does not mean that the other features become irrelevant to the identification of the degree of constitutivity, as I will clarify in the next subsection.

Empirical shielding (ES). On the one hand, the extent to which some element can be discarded and replaced, given the empirical evidence that can be obtained from within its framework, is in part dependent on its degree of abstraction. In other words, abstraction *shields* a representation from empirical testing, since abstraction itself amounts to the bracketing of empirical details. In the case of theories, once some empirical regularity, for instance, has been given a general formulation, especially by means of highly symbolic tools, this formulation, rather than being subject to empirical disconfirmation, might be viewed as inapplicable to certain types of phenomena, while applicable to others. On the other hand, the generality of the framework of inquiry, in which these representations are embedded, has an impact on the possibility of testing them empirically.[64] The broader the framework, the more it seems difficult to empirically test its constitutive components, because of the absence of alternative frameworks. Testability stands, therefore, in roughly inverse proportion to the degree of generality of the framework of inquiry, and to the degree of abstraction that *shields* an epistemic component from being empirically tested. I call the combination of these factors the degree of empirical shielding of a component, that represents the extent to which some epistemic component *cannot* be empirically tested from within the framework of inquiry in which it is embedded.[65]

[63] This is not to say that an epistemological analysis of the historical development of a scientific framework is not crucial to assess the degree of generative potential of a certain component within that framework. Quite the contrary: only historical analyses can reveal the extent to which a certain epistemic component has generative potential. That does not mean that generative potential is only determined by historical accumulation of dependence-relations. Historical accumulation can, and often does, increase generative potential, but conflating the historical aspect with the potentiality aspect hides those cases in which other factors determine the high generativity of some components.

[64] Certainly, other contextual and contingent factors may limit the possibility of testing certain components of a framework (e.g., limited technological advancements and further practical constraints).

[65] Mathematical and logical principles – one could say - are devoid of empirical content and, therefore, should be applicable in every domain or framework. However, as Wimsatt (1987: 25) points out, even in the case of mathematical and logical theorems "there seem to be antecedent conditions which must be met for the relevant

2.4 Application and advantages of the framework

2.4.1 Reframing Friedman's example of Newtonian mechanics

Now that I have introduced my main analytic features, I can start putting my account to test. First, I will show how Friedman's interpretation of the constitutive *a priori* can be deconstructed, to highlight the gradualism underlying his view. For reasons of brevity, I will only provide a reinterpretation of the example of Newtonian mechanics, but the same could be done for that of special relativity. Fig. 2.1 provides an attempt to represent graphically the example of Newtonian mechanics, introduced in section 2.2.2, in terms of my features.

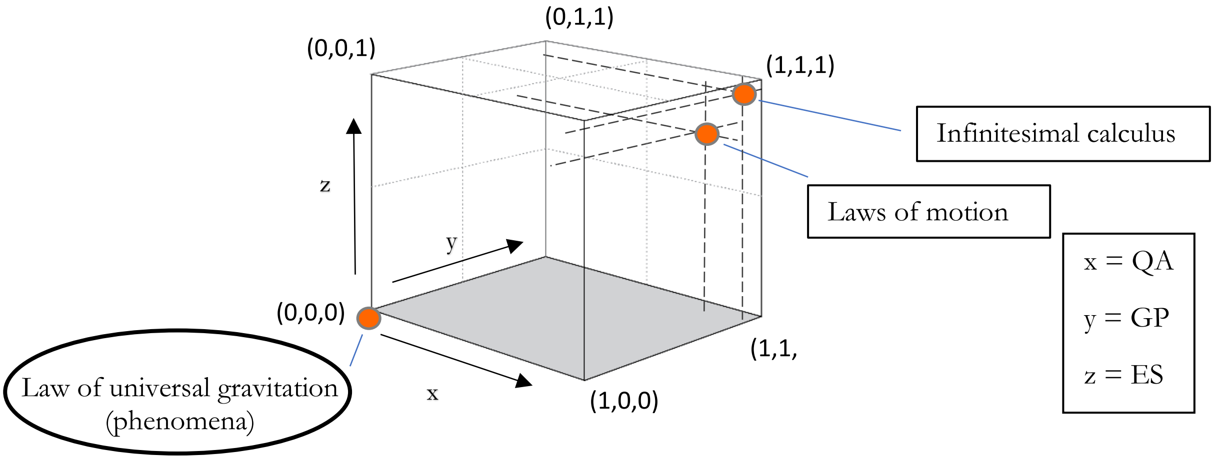


fig. 2.1: Graphic representation of the components of Friedman's (2001) analysis of Newtonian mechanics in terms of quasi-axiomaticity (QA), generative potential (GP), and empirical shielding (ES). In a circle, the domain of phenomena represented by the law of universal gravitation

Once we frame the domain of phenomena represented by the law of universal gravitation as our *explanandum*, by attributing to it value 0 for all the three features, we observe that the values of quasi-axiomaticity, generative potential, and empirical shielding for those elements that are constitutive of the law of universal gravitation in Friedman's analysis increase.[66] I must highlight

---

theorems to apply. Where these antecedent conditions are explicitly specified, they are often regarded as a part of the apparatus, [...] which does not apply everywhere". In my own terminology, albeit they have high degrees of empirical shielding, many mathematical and logical principles may have limited degrees of generative potential, thus lacking a crucial feature characterising constitutivity.

[66] The expression 'domain of phenomena' already refers to a *representation* of these phenomena, in this case by means of the law of universal gravitation. I must stress that the relationship between the law of gravitation, the laws of motion, and the calculus is partially determined by the framing of the domain of phenomena captured by the law of gravitation as our explanandum. Thus, the distance in the cube between the law of gravitation (phenomena) and the laws of motion appears as larger than what it might look like in Friedman's three-level account of theories, because we have framed the former as explanandum by

that the laws of motion are less shielded than the calculus, since a mathematical *language* cannot be disconfirmed in the same way as empirical laws, such as by means of experimental procedures directly appealing to empirical facts.[67] They also have less generative potential, because they are themselves formulated by means of conceptual structures made available by the calculus, as I pointed out in section 2.2.2. As a result, the calculus might be said to be more constitutive than the laws of motion within the context of Newtonian mechanics.

Some clarifications about the role of my three features and the relationships between them are due. Generative potential is certainly the most salient feature, since it identifies the core of the constitutive character in the general sense I outlined, that is, the capacity of some epistemic components to support conceptual dependencies with respect to more specific conceptual tools for empirical inquiry. According to Friedman, as in most of both traditional and contemporary interpretations, these elements are themselves conceptual structures, and they are often related to the workings of our scientific *theories*, but this need not be the only case. What is important here is that, by following Friedman's interpretation, only those elements count as constitutive, which approximate a *maximal* degree of generative potential within a framework. This requirement is motivated by Friedman's attempt to establish a universal and sharp dichotomy between constitutive and non-constitutive components, but, in my view, it is neither a necessary nor a desirable feature of a general account of constitutive elements. What is more, Friedman's relativised *a priori* show degrees of quasi-axiomaticity and empirical shielding that also tend to the maximum. Quasi-axiomaticity guarantees the fixity of the principles, that is the constancy of their epistemic role with respect to the constituted element. In the example of Newtonian mechanics, for instance, we have that the epistemic function of the laws of motion, i.e., that of providing the concept of inertial frame, is *always* presupposed when the law of universal gravitation is deployed for some epistemic goal (e.g., for providing explanations, making predictions, etc.) or when it is empirically tested. Finally, empirical shielding is the measure that captures the degree of shielding of some components from empirical testing and disconfirmation. As we saw, the calculus approximates a maximal empirical shielding because it is a mathematical language; the laws of motion, instead, have

attributing to it, somehow artificially, the values (0,0,0), that is, by assuming that they are not constitutive with respect to anything else, which is, of course, an idealisation. The latter move, which is essential to the gradual account, implies that we could frame a different epistemic component as our explanandum, even within the same theory.

[67] This does not mean that a mathematical language cannot be empirically tested *at all*. The expressive adequacy of a mathematical language can be evaluated based on its capacity to express an empirically adequate theory.

high empirical shielding, since they cannot be tested from within the framework of classical mechanics, but they are, in a sense, closer to empirical phenomena than the calculus. As it appears, Friedman's interpretation of the constitutive *a priori* – even though it allows some relativised *a priori* principles to have empirical content, as in the case of the laws of motion – requires maximal or close to maximal degrees of quasi-axiomaticity, generative potential, and empirical shielding. On the other hand, as I argued in section 2.2.3, it requires gradualism, which is, however, constrained by his three-level interpretation of the structure of physical theories.

2.4.2 Advantages of the new framework

As I have argued, Friedman's view does not have the resources to account for many varieties of constitutive components. Still, decoupling the feature of gradualism from his specific analysis of the structure of scientific theories provides useful insights. Epistemic components with a constitutive character might show non-maximal degrees of quasi-axiomaticity, generative potential, or empirical shielding. While Friedman's view appears to require constitutive elements to approach a maximal degree of quasi-axiomaticity, this does not seem to be required, for instance, in the historical process by which measurement procedures come to constitute the meaning of quantity terms in physical theories.[68] By allowing for degrees also in the case of the other parameters, generative potential and empirical shielding, my account makes room for a unified analysis of the epistemic components involved in coordination, including those that are *less* constitutive than the ones Friedman focuses on. Yet, a relatively high degree of generative potential will plausibly be a central feature of all elements having a role in achieving some form of coordination. Similarly, a very low degree of empirical shielding will hardly be a feature of highly constitutive elements. This is because the possibility of a certain element to be tested empirically in a variety of ways would be in contrast with it being a feature that depends significantly on the scientific framework under analysis, rather than it simply being a feature dependent on a concrete system.

In general, my account is meant to provide a tool for comparing the constitutive character of different elements *from within* a scientific framework.[69] As we saw, the precision in assigning values to the three parameters is limited and it only allows us to identify ordering relations between

[68] Padovani, personal communication.

[69] The possibility of making meaningful comparisons *across* scientific frameworks introduces several complications. For one, it would require a sufficient degree of similarity between the explananda of the different frameworks, e.g., in the case of theories dealing with (roughly) the same domain, or when both explananda have a similar form (e.g., when both are laws expressed in mathematical form, or when both are quantity concepts, etc.).

epistemic components in a framework. This limitation increases the importance of case studies to demonstrate the fruitfulness of a gradual understanding of the constitutivity of certain epistemic components based on quasi-axiomaticity, generative potential, and empirical shielding, as characteristics that come in degrees and are context sensitive. My account is not supposed to be used just to analyse what might be considered as exemplar cases, like that of Newtonian mechanics discussed by Friedman. Given a scientific framework, there will be some elements which, for any such framework, exhibit different degrees of quasi-axiomaticity, generative potential, and empirical shielding, and, therefore, their characterisation as 'constitutive' will apply to some degree.

The gradualism of my account does allow for a differentiation from Quine's entrenchment picture. Cases of, for instance, well-established empirical constants, such as Avogadro's number or Boltzmann's constant – which would have to count as highly entrenched from the point of view of Quine's web-of-belief, since their empirical adequacy has been robustly tested in many different ways – would hardly count as constitutive, given their very low empirical shielding.[70] It is difficult to map my three parameters onto Quine's view for two reasons. By overemphasising the relevance of pragmatic cost of replacement, Quine underestimates the fact that certain components are not empirically shielded simply because of the pragmatic constraints of their testing. In addition to that, the character of generativity is not highly valued in Quine's view, since he seems to assume that the elements towards the centre of the web-of-belief (mathematical and logical principles) are more entrenched, while, on the contrary, many logical and mathematical components have a high quasi-axiomaticity, but are not necessarily high in generative potential.[71] My parameters are also hidden in Friedman's view, but in a different way. In Friedman's account these parameters are conflated, in that they are correlated by definition so that, in the cases in which they are all present at a very high degree, they indicate the constitutive elements. Finally, my account diverges from Wimsatt's view, because it does not require the character of generativity to strictly correlate with the persistence, along the historical development of science, of the components with a constitutive character. Thus, it can be used also to make sense of cases of radical scientific change, which are instead well accounted for by Friedman's view.

In conclusion, the flexibility of my novel account has the double benefit of accommodating both the quasi-independence of scientific frameworks – in that it does not presuppose any hierarchical

[70] Tsou (2010) makes the case for Avogadro's number being entrenched in the Quinean sense, but not constitutive.
[71] Cf. above, footnote 65.

relationship or the absence of inconsistencies and gaps between scientific frameworks – and the diversity of constitutive components. This account, emerging from the disaggregation of Friedman's constitutive *a priori*, Quine's entrenchment, and Wimsatt's generative entrenchment, and from their reshaping as quasi-axiomaticity, generative potential, and empirical shielding, allows for a better representation of the variable landscape of scientific knowledge, resisting the Quinean image of a flat web-of-belief; it accounts for scientific change of non-revolutionary, non-theory-centric, and non-physics-centric nature, overcoming the rigidity of Friedman's perspective; and, finally, it also allows for cases of radical conceptual change that resist Wimsatt's naturalistic developmental picture.

CHAPTER 3

3.1 Constitutive elements beyond physics: a peek into the life sciences

3.1.1 Theory extension and its impact on coordination

In the previous chapter, I discussed an example from the literature on constitutive principles in science, i.e., that of Newtonian mechanics, and reframed it in terms of my own account. The transition from Newtonian mechanics to relativistic physics is often considered as a case of radical conceptual discontinuity, if not as a scientific revolution, where this is a necessary assumption of Friedman's argument that some theoretical principles are constitutive of Kuhnian paradigms. Yet, my new account for constitutivity is compatible with units of analysis other than Kuhnian paradigms, with non-revolutionary scientific change, and with sciences other than space-time physics. With this chapter, I build up to the claim that the degree and context-sensitivity of my account is useful to make sense of some episodes of theoretical change in sciences other than physics, i.e., in evolutionary biology, and which cannot be described in terms of Kuhnian paradigm shifts.

Indeed, theoretical changes are ubiquitous in scientific inquiry but only in a few cases they imply radical conceptual shifts. Theories can change piecemeal or at different levels of generality, but still these changes can have an impact on the coordination between the abstract representational tools used in the theory and the phenomena they are intended to represent. For instance, even cases often described in terms of scientific growth – in contrast with episodes of change or revolution – may lead to a transformation of the coordination between abstract theoretical resources and the phenomena that those resources are supposed to account for.

When a scientific theory T extends its range of application to phenomena beyond the ones for which it was originally conceived and developed, we clearly do not witness a case of theory replacement but, rather, of theory *extension*. With theory extension, some phenomena – whether they were previously accounted for by another theory or not – are *internalised* by theory T. By internalisation of some phenomenon P, I refer to the outcome of the process by which P is described or explained in terms of theoretical resources available within a theory T, where the latter had already been established with respect to another set of phenomena. Therefore, the internalisation of P by means of T is a result of the extension in scope of T. More specifically, it is the explanatory or representational scope of certain theoretical principles or concepts of T that is extended to a domain of phenomena larger than the one for which they were previously used. Yet, in some cases, such an extension may affect the character of the coordination between theory and phenomena. This means that the referential relationship between some phenomena and the abstract conceptual resources used to represent them may be affected by theory extension and reflected by its successful outcome, i.e., the internalisation of further phenomena. In other words, given these parallel changes at the level of phenomena (i.e., internalisation) and at the level of theory (i.e., scope extension of some theoretical resources), the question of how theory extension affects the coordination between these two levels is an open issue.

Clearly, this is an empirical question with no simple or universal answer. Some cases of theory extension may seem to fit a rather straightforward narrative, according to which scientists discover that certain already-established empirical regularities hold also in novel domains, leading to an extension of the explanatory scope of the theory. In some others, theory extension appears to involve a process by which it is rather the referential power of certain principles or concepts that gets extended.[72] In this second case it seems that the *meaning* of certain terms used in the theory changes, by leading to what Chang (2004: 150) has labelled *semantic extension*:

> I will use the phrase *semantic extension* to indicate any situation in which a concept takes on any sort of meaning in a new domain. We start with a concept with a secure net of uses giving it stable meaning in a restricted domain of circumstances. The extension of such a concept consists in giving it a secure net of uses credibly linked to the earlier net, in an adjacent domain. Semantic extension can happen in various ways: operationally, metaphysically, theoretically, or most likely in some combination of all those ways in any given case.[73]

Chang's notion of 'semantic extension' indicates a condition in which the referential scope of a term has been extended, by stretching its applicability to empirical phenomena different from the one(s) that it was initially supposed to refer to. The increased referential power of this term can be justified by means of metaphysical assumptions, by relying on a set of physical operations (such as measurement procedures), by relying on theoretical justifications, or by different combinations of the above.

I would like to stress again that the dynamics resulting in the internalisation of phenomena and in the scope extension of theoretical resources can be very complex, since both empirical evidence and non-empirical considerations (including theoretical, metaphysical, and pragmatic ones) can be used to justify theory extension at different levels of abstraction. Consequently, as I will emphasise again

[72] This idea of extension as a cognitive operation aimed at expanding the domain of applicability of a concept or lawful statement to an empirical domain different from the one of origin was used in a similar way by the logical empiricists, who emphasised the conventional aspects of this operation. Cf. Carnap (1936, especially pp. 445-6), but also Waismann (1968, pp. 118-119) on 'open-texture' and Lakatos (1976, pp. 99-100) on 'concept stretching' for similar notions.

[73] Italics in the original text.

throughout this chapter, the understanding of the impact of theory extension on coordination may benefit from a distinction between different stages of theory extension, each of them involving conceptual resources characterised by various degrees of abstraction from phenomena. Let us consider a toy example. Imagine an astronomical theory that explains several empirical regularities in one part of the universe and expresses these regularities by means of abstract mathematical laws. These lawful statements describe the behaviour of certain entities in this system, thus also providing grounds to classify these entities, some of which are referred to as 'planets'. At a later stage, scientists find that similar regularities hold in another part of the universe with similar (but not the same) entities and extend those laws to that novel empirical domain. In other words, the scope of these laws is extended, while new phenomena are internalized. This extension may seem to provide a theoretical justification for the referential extension of the term 'planet' to similar entities in the novel domain, that is, for the semantic extension of that term.

In this toy example, some lawful statements are extended to a novel domain and one of the outcomes of this operation is the semantic extension of a concept referring to a class of concrete entities. Yet, the relationship between the extension of abstract theoretical laws and that of less abstract concepts (e.g. 'planet'), whose application depends on those laws, may not be so straightforward. In some cases, the extension of lawful statements does not merely result from the incorporation of novel empirical evidence but may itself require the modification of certain abstract mathematical tools. In the toy example, the mathematical relationships between some parameters may be subject to generalisation, reformulation, or alteration. In this way, the regularities in the novel part of the universe that scientists consider as similar to those in the known part can be subsumed under the already-established theory. This operation can produce further cascade effects in the form of constraints on how certain concepts of the theory must be extended or reformulated.

In some cases, the extension of theory may fail to precisely determine the semantic extension of those terms that are 'closer' to empirical reality than the abstract laws. In the toy example, even if the semantic extension of the term 'planet' to entities in the novel domain may seem to follow from the extension of theoretical laws in the first place, further complications may emerge. For instance, some entities in the novel domain may not behave exactly as those categorised as 'planets' in the domain of origin, although the same regularities described by the mathematical laws hold, thus potentially giving rise to a gap in the coordination between the term 'planet' and its concrete referents. In the light of this complexity, the semantic extension of the term 'planet' in the toy example is better conceived as a final, mature stage of a process of theory extension, which often starts with what I call a stage of *functional* extension of some more abstract theoretical resources, that is, their extension for modelling, explanatory or, in general, theorising purposes. In the rest of this chapter, I will get back to this distinction as a useful tool to distinguish between different stages of a specific process of theory extension.

In this chapter, I focus on a pattern of theory extension in evolutionary biology that Okasha (2018) has labelled 'endogenization'. According to Okasha, endogenization is a theoretical strategy that boils down to taking some background presuppositions so far believed not to be directly accountable for in terms of natural selection theory, but rather representing its background conditions, and describe them as results of evolution by natural selection. In Okasha's view, endogenization mirrors an increase in the generality of the theory of natural selection and, more specifically, of certain principles that make up its core. My goal is to improve Okasha's epistemological take on endogenization and show that these principles are not only more general, but also more constitutive. What I show is that, with endogenization, these principles underwent a functional extension, i.e., an extension in their explanatory and representational scope, while, at the same time, some phenomena were internalised within the theory of natural selection. Yet, the scope extension of these abstract theoretical principles also changed the criteria of application of certain concepts within the context of the theory. In other words, the functional extension of the abstract principles generated the problem of re-coordinating these principles with the new extended domain of phenomena. This re-coordination requires, as its final stage, to delimit and justify the semantic extension of those more empirical concepts whose referential scope was left undefined after the functional extension of the theoretical principles.

More specifically, within the example of endogenization, I focus on how the extension of the abstract principle of heritability in natural selection theory is related to the extension of the term 'inheritance', which refers to the concrete inheritance systems responsible for the transmission and retention of phenotypic traits. What I suggest is that a stage of the process of endogenization, i.e., the functional extension of the theory and, more precisely, the extension of the principle of heritability, required a semantic extension of the term 'inheritance' without, however, straightforwardly delimiting its new referential scope. Current discussions on the meaning of 'inheritance' reflect this state, in which a fully justified semantic extension has yet to be established, while the more abstract principle of heritability has already been fruitfully extended to novel empirical domains for purposes of theorising and modelling. In this sense, I will suggest that the constitutivity of the principle of heritability increased already in the phase of functional extension, even though its extent cannot be precisely determined until the (re-)coordination between 'inheritance' and the novel empirical domain(s) is achieved, and the issue concerning the semantic extension of the term settled.

In the rest of this section, I will outline some general features of theorising in the biological sciences that are relevant to endogenization and theory extension in general. In section 3.2, I will analyse the theoretical strategy of endogenization in evolutionary biology. More specifically, I will first introduce Okasha's view and describe two cases of endogenization in detail. Then, I will present Okasha's epistemological take on this theoretical strategy, explain how I aim at improving it, and suggest that the outcome of endogenization is a special case of what I described above as internalisation. In section 3.3, I will build up to the claim that endogenization is best viewed as an increase of the constitutivity of certain core theoretical resources of the theory of evolution by natural selection, rather than simply of their generality, as Okasha suggests. To do this, I will first show that attempts at formalising the core structure of the theory of natural selection fostered endogenization, whose second and crucial stage is an extension of the theoretical role of the core principles of natural selection (functional extension). Then, I will suggest that current discussions concerning the *meaning* of a term central to the theory of natural selection, i.e., 'inheritance', express the attempt at re-coordinating one of these core principles, i.e., 'heritability', with the novel empirical domains to which its epistemic *function* had already been extended. Therefore, I will argue that, in endogenization, the principles were first functionally extended, while the semantic extension of some theoretical terms represents the final stage of the endogenization process. Finally, I will suggest that the degree of constitutivity of these principles, expressed by the features of quasi-axiomaticity (QA), generative potential (GP), and empirical shielding (ES), increased with the use of endogenization as a strategy of theorising. In this sense, the principles became more constitutive already with their functional extension, but it is only when the semantic extension of the concepts they relate to is accomplished that the increase in constitutivity can be assessed more precisely.

3.1.2 Styles of theorising and theory structure in evolutionary biology

Evolutionary biology is often presented as a "pluralistic explanatory and predictive enterprise" (Winther et al. 2013: 976), and its current theoretical standpoints are the product of a long process characterised by relatively strong division of labour. In general, evolutionary biology makes use of a great variety of models that aim at representing different target phenomena. Model users usually restrict their focus to the variables needed to account for the phenomenon of interest by means of varying degrees of abstraction and idealisation (Godfrey-Smith 2009a). In certain cases, it is possible to integrate these models in a rather unified and cohesive fashion, but sometimes they can be incompatible with one another (Morrison 2011). Consequently, the internal relationships between models and structural components of theories in evolutionary biology is not a simple matter.

Among the many aspects of this complexity, two stand out with respect to my focus in this chapter. On the one hand, evolutionary biology often combines many kinds of theorising, that span from the highly mathematised or 'formal' (Thompson 2007) and, for some, even close to logico-deductive schemes (Tennant 2014), to the induction-based and 'material' representation of theoretical content (Love 2013). However, this variety of theoretical tools cannot simply be dismissed as a potentially surmountable lack of methodological unity. Rather, the consensus among most scientists and philosophers of the life sciences seems to lie on the acceptance that "different biological fields provide not only different theories about the same sets of phenomena, but also different *types* of theoretical results, ranging from the mostly descriptive knowledge gathered by experimental or field biologists to the largely speculative claims pursued by theoretical biologists through mathematical reasoning" (Leonelli 2009: 191).[74] On the other hand, the widespread assumption that evolutionary biology is a unitary field with a cohesive framework in the background (cf. Scheiner 2010) seems to motivate the issue of choosing *the* best single structural container for evolutionary theory, particularly in the context of the current debate opposing the Modern Synthesis and different versions of an extended evolutionary synthesis (Laland et al. 2014, Pigliucci & Müller 2010).

In the rest of this section, I will rehearse some passages of the history of evolutionary biology that are relevant to the debate on the current status of the theory of evolution by natural selection. My aim is not that of giving a simplified sketch of the history of this theory or of the open debates on its status. Rather, my focus is on some macro-level epistemological patterns that are relevant to my discussion of constitutive elements in science. For this purpose, it suffices to distinguish between three cardinal moments in the development of the theory of natural selection. The first moment is in 1859, when Charles Darwin's *The Origin of Species* made its public appearance as the first published text presenting the theory of evolution by natural selection.[75] Darwin's book was the result of more than two decades of work, which incorporated a wealth of empirical observations that he collected starting from his five-year journey on the *Beagle*. According to Darwin, nature selects individuals which possess traits that make them more likely to survive and reproduce than others. Therefore, natural selection only operates on those traits that can be passed on to the individual's offspring via reproduction. Famously,

[74] Italics in the original text.

[75] For a classic account of Darwin's path to the development of the theory of natural selection, see Ospovat (1995). For a comprehensive history of the relationships between the different evolutionary theories, see Bowler (1989).

Darwin's theory in 1859 did not account for the mechanism of heredity of traits, which eventually started being identified with the rediscovery of Mendel's work in 1900. The unification of Darwin's theory of natural selection and Mendelian genetics was a long process, rich in controversies, which required the contribution of numerous scientists with different specialisations (Mayr & Provine 1988). The Modern Synthesis or neo-Darwinian framework resulting from this effort was finally accomplished with a series of published works starting from the late 1930s and culminating with J. Huxley's *Evolution: The Modern Synthesis* in 1942, which I take conventionally as the second crucial moment of this hyper-condensed historical summary. Forty years later, Mayr (1982: 120) could still appropriately claim that the "synthetic theory of evolution is the paradigm of today". Yet, already by the 1970s, biologists from different subdisciplines had started challenging the adequacy of Neo-Darwinism as a theoretical framework for all evolutionary phenomena in the biological realm. Scientists working in fields that were underrepresented at the times of the building of the Modern Synthesis, such as embryology, developmental biology, and ecology put pressure in the direction of shifting from the Modern Synthesis to an extended evolutionary synthesis, which would include a wider range of phenomena and theoretical resources. As I already mentioned, this debate is still heated today, and I take the current state of theorising in evolutionary biology to be the third historical point of reference to consider.

3.2 Growth or change? The case of endogenization in evolutionary biology

3.2.1 What is endogenization?

While doing evolutionary modelling, biologists treat certain variables as representing the evolutionary forces deemed to be active and, therefore, as endogenous, while some other variables are assumed as background conditions, that is, as exogenous (Okasha 2018). According to Okasha, "a broad trend in evolutionary biology over many years" can be identified, in that elements traditionally belonging to the background of the core theoretical framework of evolution by natural selection have subsequently been 'endogenized' in the domain of possible evolutionary theorising.[76] In a nutshell, endogenization boils down to taking some phenomena so far believed not to be directly accountable for in terms of natural selection theory but, rather, idealised away as variables that constitute its background conditions, and describe them as results of evolution by natural selection. This theoretical strategy is so pervasive in evolutionary biology, that the latter's historical development can be viewed as a

[76] Personal communication.

"successive endogenization of variables that previous theorists had treated as exogenous" (Okasha 2018: 3).

One particularly striking aspect of this process is that the same theoretical resources deployed at later stages to account for these background presuppositions *required* – at a previous stage of development of evolutionary biology – the backgrounding of those very presuppositions, in order to be identified and developed in the first place. I believe that this aspect can be reframed as follows. Let us say that, at stage t_1 of theoretical development, certain variables (x_1, x_2, x_3) had to be fixed as background presuppositions, so that the other variables (y_1, y_2, y_3) could be isolated as those that could provide an explanation of phenomena (P_1, P_2, P_3, P_4, P_5). This means that phenomena (P_1,…, P_5) could be modelled in terms of variables (y_1, y_2, y_3), only when conditions expressed by variables (x_1, x_2, x_3) were assumed as independent from the relationship between (P_1,…, P_5) and (y_1, y_2, y_3), that is, as exogenous. However, at stage t_2, variables (y_1, y_2, y_3) are extended, so that they can model also phenomena (P_6, P_7, P_8). Phenomena (P_6, P_7, P_8), however, were previously described as background conditions in terms of those very variables (x_1, x_2, x_3) that, at t_1, were idealised away so that phenomena (P_1,…, P_5) could be successfully modelled in terms of variables (y_1, y_2, y_3). What happens at t_2, then, can be described as follows: (y_1, y_2, y_3) become *endogenizers,* because they have a role in the transformation of (x_1, x_2, x_3) – the *endogenized* – from idealised background presuppositions into phenomena that can be described and modelled in terms of (y_1, y_2, y_3). In the following sub-sections, I describe two cases of endogenization that will be helpful to substantiate the abstract characterisation of this process that I just outlined.

3.2.2 Endogenizing individuality

The Modern Synthesis achieved the integration between Darwin's theory of evolution by natural selection and the Mendelian theory of genetic inheritance. The latter, rediscovered in 1900 after almost four decades of oblivion, provided the mechanism responsible for the heritability of phenotypic traits that Darwin was unable to identify.[77] This systematic integration was accomplished under the

[77] It is well known that Darwin's *Origin of Species* did not provide a specific mechanism of heredity that could explain the transmission of characteristics from parent to offspring. In 1868 Darwin suggested the 'hypotheses of pangenesis', according to which parents transmit their characters to their offspring by means of small gemmules produced by each part of the parent organism (more on this in chapter 5, section 5.2.2). Interestingly (cf. below, section 3.2.3), Darwin (1881) attributed an important role to the environment in the mechanism of inheritance. Modifications of any part of the organism due to environmental causes could result also in modifications of the gemmules which, in turn, could be inherited by the offspring that might show the resulting changes. According to Darwin, such a mechanism would leave more room for individual differences in creating enough variation for natural selection to operate. Cf. Provine (1971, ch. 1).

assumption that natural selection operates on Mendelian populations, that is, populations of individual organisms, where these individuals are considered as discrete and genetically homogeneous units.

The assumption that natural selection operates only on individual organisms is, however, quite problematic, if considered in the context of the so-called 'levels of selection issue' in evolutionary biology. This issue relates to the fact that the biological realm is usually conceptualised as hierarchically organised in different levels (molecules, cells, organisms, populations, species). Does natural selection act only on one of these levels, or does it involve more levels of selection simultaneously? This question was at the centre of a debate that originated already at the times of Darwin but reached a turning point when, starting from the 1960s, many evolutionary biologists argued that natural selection can indeed operate on multiple levels: genes, organisms, or groups.[78] What was the impact of this argument on the assumption of individuals as discrete and genetically homogeneous units in the Modern Synthesis?

Buss (1987) emphasises that the influence of this assumption in the development of the Modern Synthesis had its origins in the work of the embryologist August Weissman at the end of the 19th century. Weissman famously argued that somatic cells – that is, cells belonging to the organism tissues and not destined to the reproductive cycle – could not be heritable because terminal somatic differentiation rules out the opportunity for a cell lineage to contribute to subsequent generations. In other words, once a lineage of cells develops certain specific functional traits required by its specialised role within an individual organism (i.e., it undergoes differentiation), these cells cannot participate in the transmission of individual traits anymore. Even though violations to Weismann's theory were already known, embryological concerns did not have a central role in the intellectual effort leading to the Modern Synthesis, since the biologists involved in this enterprise all worked on organisms which followed Weissman's rule.[79] According to Buss (1987: 8), this fact is the historical root of the tacit incorporation of the axiom that individuals are the sole unit of selection in the neo-Darwinian framework: "If not all cells contain heritable material, selection is necessarily on the individual, not on the cells or their constituents".

[78] Cf. Hamilton (1963), Lewontin (1970), Price (1970, 1972), and Williams (1966).
[79] The main experimental genetic discoveries in chromosomal theory that preceded the Modern Synthesis were made by using *Drosophila melanogaster* as a model organism, most famously by T. H. Morgan's research group. The scientists (geneticists, mathematicians, naturalists) who developed the synthesis were working on dipterans and vertebrates. Fruit flies, mice, and humans are organisms where the somatic differentiation of cell lineages happens very early.

A theory that assumes the existence of individuals as a primitive axiom cannot explain the origin of individuality itself. The fact that the presence of individual organisms should not be viewed as an unquestionable background assumption of the neo-Darwinian framework is the basis for Buss's treatment of individuality as an evolved trait. Indeed, Buss notices that the lack of differentiation between somatic and reproductive cell lineages is ubiquitous in multi-cellular individuals in the Kingdom Plantae and in the Kingdom Fungi, but it is not an anomaly even in the Kingdom Animalia. This consideration, of course, is in contrast with the assumption of 'Weissmanian' individuals as the rule in the biological realm: "The ideal of the individual as an entity that may be treated as genetically uniform is at best an approximation. It is apparent that individuality is a derived character, approximated closely only in certain taxa" (Buss 1987: 20).

If individuals were the sole unit of selection, the mechanisms of cellular differentiation that suppress the possibility for certain cell lineages to reproduce, since this would be detrimental for the development of the individual organism, would be present in all individual organisms in all biological taxa. However, this is clearly not the case and it was even less the case in earlier geological eras. According to Buss, individuality originated in the Precambrian era as a result of a transition in the units of selection from single-celled organisms to multi-cellular individuals characterised by internal division of labour. This transition was prompted by a change in selection pressures on the lower level (cells). While single-celled organisms were previously undergoing selective pressures coming only from the external environment, when cell lineages started cooperating and self-organising in multicellular organisms, they started receiving selective pressures also from the internal somatic environment. In fact, the division of labour determined that some cell lineages would not undergo differentiation, since they would be destined to reproduction, whereas others would. Differentiation originated as a mechanism of suppression of the potentially detrimental effect of uncontrolled reproduction of the lower level (cell lineages), which would highjack the stability of the higher level (individual organism).

Therefore, the existence of individuals is *itself* an evolutionary product, due to the selection of higher levels on lower levels or, more in general, by synergisms and conflicts between the units of selection. Accounting for the origins of individuality in terms of natural selection theory, as Buss did, had been made possible by those biologists who – in order to argue for group-level selection in the context of the level of selection issue mentioned above[80] – had shown that even the assumption of hierarchical organisation of the biological realm could not be considered merely as a given background presupposition of the theory of natural selection. However, Buss's own analysis stimulated a lot of

[80] Cf. also Maynard-Smith (1976).

work on the "evolutionary transitions in individuality", which seeks to account for the origins of hierarchical organisation in Darwinian terms.[81] The specific result of Buss's and other contributions can be understood in epistemological terms as the endogenization of the axiom that individuals are homogeneous and discrete genetic units, within the theoretical scope of evolution by natural selection. Individuality, which was an exogenous variable during the theoretical effort leading to the Modern Synthesis, was later endogenized and explained by evolutionary biologists in terms of the same theoretical resources whose development required the backgrounding of individuality in that previous stage of theorising.

3.2.3 Endogenizing the environment

Within the neo-Darwinian framework, the environment is considered as the factor that exerts selective pressures on organisms. Organisms, in response to these pressures, evolve adaptations via natural selection in order to 'fit' the environment. Therefore, the environment is treated as an exogenous variable, it is assumed as an external factor with respect to the workings of evolution by natural selection.

This assumption has been challenged by researchers working in the field of niche construction theory. These scientists emphasise that organisms can fit the environment not only by evolving adaptations through natural selection, but also by transforming the environment in the course of their lifespan. They do so, for instance, by regularly modifying local resource distributions, choosing and changing habitats, or constructing artefacts. These activities are widespread in the animal kingdom, among both vertebrates and invertebrates, from beavers building dams and badgers constructing burrow systems, to birds building nests, and yet to spiders weaving nets, and ants and termites building nests. In addition, it is typical of plants, too, since they change numerous traits of their habitats, including temperature, humidity, chemical composition, acidity, patterns of light and shade, etc. Most organisms contribute to the construction of their environmental niches.

In an influential book, Odling-Smee at al. (2003) argue that these niche constructing activities emerge as responses to external selective pressures. At the same time, by transforming the environment, these activities contribute to changing the selective pressures which, in turn, prompt further adaptive

[81] For early works, see Maynard Smith & Szathmáry (1995) and Michod (2000). See Okasha (2005) for a historical perspective on the units of selection debate.

responses from the organisms and, therefore, affect their fitness.[82] In other words, niche construction is not simply a modification of the environment *per se*: it is a modification of the *selective* environment induced by organisms themselves. As a result of this change in selection pressures, members of many species inherit the cumulated environmental changes that previous generations have induced. The effects of niche construction can overturn external sources of selective pressure and give rise to unusual evolutionary trajectories and equilibria by, for instance, fixating otherwise deleterious alleles.

The interactions between genetic and ecological inheritance can echo into macroevolutionary patterns, as in the case of the evolution of photosynthesis in early bacteria, which led to an increase of oxygen in the atmosphere and, consequently, to the evolution of organisms capable of aerobic respiration (Erwin 2008, Danchin et al. 2011). These effects are particularly striking in the case of human evolution. Laland et al. (2010) summarise decades of studies on how gene-culture interactions have shaped human evolution. As they highlight, anthropological studies and data from the human genome systematically converge in showing that numerous genes have been subject to positive selection as a response to cultural practices.[83] Therefore, human evolution cannot be simply understood as a process of adaptation to changes in the environment caused by events beyond human control. On the contrary, gene-culture dynamics can have fast and sizeable evolutionary consequences, so that it can be considered one of the most relevant patterns of human evolution.

In sum, according to niche construction theorists, the organism-environment interactions typical of a species cannot be ignored when analysing evolutionary dynamics, since they fundamentally affect the strength and direction of the selective pressures on that species and others. Including the process of niche construction in the domain of phenomena accounted for by natural selection allows for the incorporation of environmental components and their interactions with the biological realm in evolutionary modelling. In epistemological terms, this move is prompted by recognising the idealising character of the assumption that the only way for organisms to respond to selection pressures from the environment is by evolving adaptations. This idealisation is not universally justifiable, since niche construction demonstrates how organisms can fit the environment by transforming the latter, and this variable should be taken into consideration in evolutionary modelling, when relevant to the modelling goal. Incorporating niche construction represents an attempt at endogenizing the environment, that

[82] Earlier papers outlining the core ideas of the book are Laland et al. (1996, 1999), Odling-Smee (1988), and Odling-Smee et al. (1996). Lewontin (1982, 1983) anticipated some lines of research they developed, whose importance was already clear to Darwin (1881). Cf. above, footnote 77.

[83] One classic case study is that of how dairy farming provided the selective environment for the positive selection of alleles for lactose tolerance in adults. Aoki (1986) and Feldman & Cavalli-Sforza (1989) developed the first models. See also Holden & Mace (1997) and Gerbault et al. (2011).

is, at bringing the formerly exogenous variable of the environment within the scope of evolutionary theorising. What is endogenized, in this case, is the *selective* environment, which can be understood as a variable co-evolving with the traits (i.e., the niche constructing activities) evolved by the organism to interact with the environment itself.

3.2.4 Endogenization results in a special case of internalisation

The two cases of endogenization that I presented are just a representative sample of how this theorising strategy works. Other variables that were taken as background conditions of evolution by natural selection, including mutation rate, sex, genetic recombination, fair meiosis, population structure, etc., have been subject to endogenization. This process did not unfold exactly in the same way in all cases, and it did not bring identical results. For instance, some variables, e.g., the environment, have only been partially endogenized, whereas others have been completely explained away, as in the case of mutation rate (cf. Okasha 2018). However, all these lines of theorising have a common underlying epistemic mechanism. By means of the notion of endogenization, Okasha captures a specific type of theory extension in evolutionary biology, one by which phenomena that were idealised away as background presuppositions by previous formulations of Darwinian evolution could later be accounted for by that theory.

On a surface level, the progressive endogenization of background variables in the history of evolutionary theory fits a narrative of cumulative progress, according to which "evolutionary biologists have responded to various crises by augmenting the preexisting framework, building on what was already there" (Pigliucci 2007: 2743). In other words, Okasha's notion of endogenization seems fully consistent with the view that evolutionary biologists react to the discovery of new phenomena or to the emergence of conceptual issues by extending the application of already-established theoretical tools, rather than by developing alternative theoretical frameworks or suggesting radical conceptual shifts. In addition, his analysis of endogenization complements previous scholarly work focusing on the strategy of extending evolutionary theorising beyond the life sciences, as it was used in the study of animal behaviour (Avital & Jablonka 2000, Jablonka & Lamb 2005), in certain areas of psychology (Campbell 1960), and in the study of culture, including language, morality, and science itself (Boyd & Richerson 1985, Cavalli-Sforza & Feldman 1981, Hull 1988). By showing how endogenization worked as a strategy of theory extension directed at certain presuppositions of the theory itself, Okasha provides further systematic evidence of the historical process by which the theoretical resources of evolution by natural selection progressively expanded their scope.

From an epistemological point of view, Okasha (2018: 2) characterises endogenization as "a particular way in which the generality of evolutionary theory has been increased over time". More specifically, he seems to suggest that the generality of the core principles of Darwinian evolution, i.e., the possibility to provide an abstract formulation of these principles, enabled endogenization as a strategy of theorising to flourish.[84] In the terminology that I introduced in chapter 2, this means that the empirical shielding (ES) of the core principles of Darwinian evolution, i.e., the extent to which they are shielded from empirical testing due to both their generality and their degree of abstraction, was the crucial feature underpinning endogenization which, in turn, led to an increase in generality of natural selection theory. In my view, the empirical shielding of these theoretical principles was certainly crucial to endogenization, especially because it increased over time, as I will emphasise in the next section. Yet, I will suggest that the two other features that I introduced in chapter 2 – quasi-axiomaticity (QA) and generative potential (GP) – also increased over time and played a central role in the process of endogenization. Therefore, I will suggest that it is not just the generality of these theoretical principles that increased with endogenization, but also their constitutivity. The first step towards this claim is to understand endogenization as a specific sort of theory extension that, on the one hand, results in the internalisation of some phenomena and, on the other, has an impact on the coordination between the theoretical principles and the phenomena they are supposed to represent.

Wimsatt (1987) focuses on evolutionary theorising beyond evolutionary biology to understand how theory extension works. According to Wimsatt, in the case he analyses certain theoretical resources were abstracted away from their context of origin, so that they become less tied to the empirical details of the domain within which they were initially formulated. This process of abstraction makes them 'portable' from the old domain to a new one, in which they can serve some explanatory or representational role. This is the case, for instance, of evolutionary epistemology, which describes the development of science according to the explanatory scheme of evolution by natural selection.[85] Wimsatt's analysis focuses rather extensively on how the character of portability of those conceptual resources and the possibility to extend their scope come about through abstraction, rather than on the pragmatics of their actual application in a new domain.

[84] Okasha (2018: 18) carefully emphasises that endogenization shows that the core Darwinian principles do not "bear the explanatory burden in evolutionary biology, but rather those principles as they operate in specific biological settings, in the presence of additional contingent biological features", thus rejecting the possibility of a radical form of reduction of all evolutionary explanations in terms of these core theoretical resources.
[85] Cf. above, chapter 2, section 2.3.1.

Yet, it seems to me plausible to say that this epistemic process can conclude successfully only by 'rewiring' – so to speak – these theoretical tools in the new domain of application. In other words, some novel empirical phenomena can justifiably be described or explained in terms of theoretical tools already developed in a different framework only once these phenomena are coordinated with the conceptual resources abstracted away from the domain of origin. Eventually, the novel empirical phenomena undergo what I called internalisation, in the sense that they are brought within the scope of a certain theoretical or explanatory strategy. However, the phenomena of the new domain are *successfully* internalised when coordination between the concepts from the old domain and the phenomena from the new domain is achieved, that is, when the extension of their referential scope is justified. As in the toy example of section 3.1.1, the functional extension of some laws or theoretical principles to a novel domain may constrain the conditions of applicability of certain terms (e.g., 'planet') without entirely specifying the precise limits of their semantic extension in the extended domain of phenomena.

In this sense, in the cases of endogenization that I discussed above scientists seemed to achieve the internalisation of some phenomena, by applying certain theoretical resources – the endogenizers – to empirical domains that were previously not accountable for in terms of natural selection theory. Yet, there is a peculiar feature of endogenization, with respect to other dynamics of theory extension that achieve internalisation. The extension of applicability regards phenomena which, at previous stages of theorising, were captured by variables (i.e., environment, individuality, hierarchical organisation, etc.) that represented certain background conditions for the applicability of the theory of natural selection itself. They were assumed as fixed, so that the theory of natural selection could explain another host of phenomena, such as adaptation, speciation, and phylogenesis. At a later stage, however, the very same theory supplied an explanation of these variables, which from fixed idealised presuppositions became internalised phenomena. If we assume that endogenization results in a special case of internalisation, we must consider its unfolding through different stages, in a similar way to what Wimsatt suggests, where its final stage is the achievement of a successful form of coordination between theoretical resources and the extended domain of phenomena.

Certainly, contingent historical factors come in the way, once one tries to trace the development of endogenization from the abstraction of theoretical principles, to their (functional) extension, and yet to their coordination with the extended domain of phenomena, which provides a justification for the semantic extension of certain concepts of the theory. As I have mentioned, in the case of individuality endogenization came to dispel the effects of an assumption – that of individuals as discrete and genetically homogeneous units – which was warranted an unjustified role during the development of the Modern Synthesis. Clearly, it was a matter of historical contingency that the scientists working at the Modern Synthesis took the Weissmanian notion of individual for granted, although plausibly there were pragmatic concerns guiding this choice. After all, any scientific inquiry requires holding some element fixed, to test or theoretically develop others.

However, the fact that different lines of theorising in evolutionary biology converged on the very same epistemic pattern, that is, on endogenization, requires a deeper explanation. In the next section, I suggest that the strategy of endogenization was fostered by the formalisation and achieved through the functional extension of the core principles of natural selection which, in turn, have to be (re-)coordinated with the extended domain of phenomena, so as to justify the semantic extension of certain core concepts of the theory. This claim eventually bears on my discussion of constitutive elements in science, because it enables to understand the epistemic role of these core principles in terms of empirical shielding, quasi-axiomaticity, and generative potential.

3.3 The epistemic history of the core Darwinian principles

3.3.1 <u>Formalising the core Darwinian principles</u>

So far, I showed that scientists achieved, through endogenization, a peculiar form of internalisation, in the attempt to explain certain background variables of natural selection theory by recurring to theoretical and conceptual resources of that same theory. Therefore, these theoretical and conceptual resources were necessary for endogenization to be carried out. The explicit identification and 'extraction' of those resources from the theoretical structure of origin is unlikely to be a straightforward matter, and it may not be a necessary condition to achieve internalisation in general. Yet, formalisation is one explicit way to identify and extract the theoretical and conceptual core of a theory. In this sub-section, I suggest that the formalisation of the core Darwinian principles was a first crucial step of the scope extension of theoretical resources that was central to the process of endogenization.

According to the standard philosophical understanding, a formalisation of a scientific theory is "an abstract representation of the theory expressed in a formal deductive framework" for which there is a complete specification of 1) what constitutes a well-formed formula, 2) all the permissible rules of inference, 3) a set of primitive formulae of the theory, and 4) an interpretation (semantics) of the formalism (Thompson 2007: 485-486). However, formalisation represents a broader category than mathematization, especially in the context of the biological sciences, and it has been characterised as the "drawing of a form/content distinction so as to make possible the study of form independent of content" (Griesemer 2013: 302), where this distinction can come in different varieties and degrees, and for different purposes. By means, for example, of abstraction, i.e., the bracketing of empirical details, the empirical content of a theoretical principle or set of principles can be 'stripped out' from their formulation, so that the identification of the regularities captured by those principles is less constrained by the specific details of the empirical domain of origin.[86] Formalisations can resort to abstraction together with certain forms of idealisation, which often consist in isolating certain variables involved in the explanation of a phenomenon and imagining away other variables so as to simplify the representation of the phenomenon under investigation.[87]

In the context of the debate on the levels of selection that I mentioned above, Lewontin (1970) formulated a distilled version of Darwin's (1859) fundamental explanatory structure in the form of three principles that underpin the notion of evolution by natural selection. These three principles can be applied to any of the different units of selection (molecule, cell, organism, group of organisms, species), and, therefore, have been considered as a sort of 'recipe' for evolutionary change (Godfrey-Smith 2009b). The principles, as worded by Lewontin (1970: 1), are the following:

i. Different individuals in a population have different morphologies, physiologies, and behaviors (phenotypic variation)
ii. Different phenotypes have different rates of survival and reproduction in different environments (differential fitness)
iii. There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable).

What Lewontin provides is a formalisation of evolution by natural selection in terms of the principles of variation (V), differential fitness (DF), and heritability (H). On the one hand, this is the result of an individuation and extraction of the three formal principles from the informal Darwinian formulation of the theory (Griesemer 2013: 306-7). This extraction is underpinned by an operation of abstraction from a large amount of empirical observations from which Darwin originally inferred the structure of the mechanism of evolution by natural selection. The three core principles evidently involve abstract mathematical and statistical concepts (variation as a statistical distribution, fitness as rates of survival

[86] Abstraction can come in different varieties, such as reification, hypostatisation, singularisation, concretisation, individualisation, analogy-making, etc.
[87] Idealisation is often described as the omission and/or distortion of relevant variables with the aims of simplification. Still, as in the case of abstraction, it is a placeholder term that refers to a variety of activities. For a survey of different modes of idealisations deployed in the sciences, see Auyang (1999). For more on abstraction and idealisation in evolutionary biology specifically, see Godfrey-Smith (2009a).

and reproduction, heritability as intergenerational correlation), even though they are not expressed in mathematical form.[88] On the other, this formalisation, clearly resorts to idealisation, since it leaves out a number of variables deemed as crucial by Darwin himself, such as the famous notion of struggle for existence.

Historically, Lewontin's formalisation was aimed at solving the issue of levels of selection and, although it was perhaps the most influential, it originated in previous attempts at providing a rigorous formulation of the theory of evolution by natural selection. As Barberousse and Samadi (2015) point out, Lewontin was influenced by Mary Williams' (1970, 1973) axiomatisation, that was circulating already at the end of the 1960s. Williams's explicit goal was to "provide a non-ambiguous and non-tautologous translation of 'survival of the fittest'" (Williams 1973: 84) against charges of circularity of natural selection theory, and in the attempt to improve the naiveties of previous formal treatments.[89] In addition, Lewontin's was not the last attempt at formalising the theory of natural selection. Among more recent attempts at formalisation, I only mention Godfrey-Smith (2009b), who provides a grid with degree-sensitive parameters for a more inclusive conceptualisation of evolution by natural selection, and Barberousse and Samadi's (2015) formalisation in terms of probabilistic laws, which aims at mirroring the contemporary practices of evolutionary biologists. Both formalisations add some variables to the original three Lewontinian principles, thus de-idealising from it, and problematise the correct interpretation and formal characterisation of the three principles themselves.

Indeed, the formalisation of a theory can be pursued in order to achieve several different goals. One of them is to uncover the fundamental assumptions of a theory and the relationships between them. When these assumptions receive a rigorous formulation, it is easier to derive their theoretical implications and test them against empirical evidence. This, in turn, allows for the testing of certain epistemic virtues of the theory, such as internal consistency and explanatory power. Rather independently of the goal, formalisation has a peculiar impact on certain theoretical components that

[88] See Lewontin (1985: 76) for a slightly different, and somewhat clearer, formulation of the core principles. More specifically, Lewontin (1985) expresses the principle of heritability more explicitly in terms of intergenerational correlation of variation, rather than involving the notion of fitness, as in the 1970 paper. Cf. Godfrey-Smith (2009b) for a discussion.

[89] Williams provides a formalisation of natural selection by means of seven fundamental axioms expressed in set-theoretic terms. On these grounds, she rejects the two assumptions behind the charges of emptiness against Darwinian evolution, i.e., the fact that all words used in a scientific theory should be defined, and the assumption of a single fundamental level at which the 'primitive' entities of a theory should belong and of a single, equally fundamental, class of entities to which the theory should apply. Williams' formalisation was subject to many criticisms. For a summary of reactions to Williams' work and a critical perspective, see Jongeling (1985).

can be understood in terms of two features that I introduced in chapter 2. In my account, by 'empirical shielding' (ES) I referred to the extent to which some epistemic components *cannot* be empirically tested from within their own framework of inquiry. By means of abstraction, that is, the leaving out of empirical details, which occurs during the formalisation of an epistemic component, the empirical shielding of a theoretical principle or concept increases. For instance, take the following two statements: (1) "different phenotypes have different rates of survival and reproduction in different environments", and (2) "more individuals with phenotype *A* than with phenotype *B* in a population F survive under the environmental conditions x, y, z". It is harder to produce empirical evidence to disconfirm (1) since its generality and abstractness partly shield it from empirical testing (e.g., what counts as a 'different environment' may widely vary depending on context and purpose of the modeller). The process of abstraction enables the extraction of certain concepts or principles from the specific empirical context in which they originated and their use in other empirical domains by decreasing the possibility to disconfirm them via empirical observation or testing.

By 'quasi-axiomaticity' (QA) I referred to the extent to which some epistemic component is a fixed element within a framework of inquiry, with respect to its epistemic *function*, so that the other elements can be developed or tested. For example, the more a concept (e.g., 'fitness') is rigorously defined (in this case, as a measure of reproductive success), the more it is possible to deploy it unambiguously for the function of quantifying the effects of a phenomenon (e.g., to test the effects of selection) across a wide variety of empirical inquiries. This does not mean that formalisation is the only way to set fixed some epistemic components, nor that quantification is the only epistemic function that quasi-axiomaticity relates to. This short example is just meant to clarify that the fixing of the epistemic function of a component (via formalisation or other means) endows that component with a quasi-definitional status, thus (temporarily) bracketing the possibility to modify or replace it, so that other parts of the theory comprising it can be developed or empirical claims based on it tested. By means of idealisation, formalisation can fix the epistemic function of some theoretical components of a theory via a rigorous formulation that posits them as the most important variables, or it can isolate some other variables in the background, in both cases increasing their quasi-axiomaticity. The fixing of the epistemic function must not be confused with the fixing of the *meaning* of a concept or principle. By function, I mean the role that a certain epistemic component fulfils within a certain scientific framework[90]. I will go back to this distinction between function and meaning of an epistemic

[90] For my understanding of what a scientific framework is, see above, chapter 1, section 1.3.3.

component when analysing in detail the status of one of the core Darwinian principles in an endogenized domain in section 3.3.3, i.e., the principle of heritability in niche construction theory.

As the case of the three Lewontinian principles of variation, differential fitness, and heritability shows, formalisation contributed to the identification and extraction of certain fundamental theoretical resources from the general structure of the theory of natural selection, but it also had an impact on their conceptualisation, by fixing their epistemic function as enabling conditions for natural selection to occur. By means of idealisation, they were selected as the most relevant variables of the theory, i.e., those to be held fixed as presupposed in further empirical investigation. This aspect can be understood as an increase of their quasi-axiomaticity within the domain of evolutionary biology. On the other hand, they received, by means of abstraction, a formulation independent of empirical details from which they originated and, therefore, their empirical shielding increased as a result of the increase in their generality.

3.3.2 From formalisation to internalisation through functional extension

So far, I have argued that formalisation had an impact on the quasi-axiomaticity and empirical shielding of the core Darwinian principles. In this subsection, I make two further steps. On historical grounds, I suggest that Lewontin's formalisation underpinned the endogenization of hierarchical organisation and, in turn, of individuality. This step gives me grounds to argue that, from an epistemological point of view, attempts at formalisation could be viewed as the first stage of endogenization in evolutionary biology. This argument is based on the recognition that increases of quasi-axiomaticity and empirical shielding often lead to increases of generative potential, that is, the capacity of a theoretical component to sustain large amounts of conceptual dependencies.

As I explained above, by means of idealisation, formalisation fixes the epistemic function of certain epistemic components, thus increasing their quasi-axiomaticity. In the case I am analysing, the core Darwinian principles are set fixed as quasi-axiomatic theoretical components that function as modelling tools for certain phenomena, so that these can be represented as products of natural selection. At the same time, by means of abstraction, formalisation detaches those components from the specific empirical content of their original domain of application, therefore increasing their empirical shielding. Consequently, the resulting idealised formal principles can be applied to a wider range of phenomena (functional extension) and, therefore, can sustain a larger amount of conceptual dependencies (higher generative potential), in virtue of the fact that their epistemic function has been fixed (higher quasi-axiomaticity) while their connection with empirical phenomena is more abstract (higher empirical shielding), in comparison to previous, less formal formulations. The second stage of endogenization is the functional extension of the core Darwinian principles, such that they are used to model phenomena formerly represented as idealised background variables (fig. 3.1). As an upshot of functional extension, the principles of variation, differential fitness, and heritability have a higher generative potential and the formerly idealised background variables become internalised phenomena.[91]

Stage	1) Formalisation	2) Functional extension
Outcome	Abstraction of the principles → ⇧generality Idealisation → fixing of their epistemic function	Use of the principles to model previously backgrounded phenomena → internalisation of phenomena
Parameters	⇧empirical shielding ⇧quasi-axiomaticity	⇧generative potential

fig. 3.1: Table summarising the outcomes of the process of endogenization from formalisation to functional extension, and their effects on the empirical shielding, quasi-axiomaticity, and generative potential of the core Darwinian principles

Let us go back to Lewontin's formalisation. His formulation of the three principles of variation, differential fitness, and heritability clearly shows relevant differences from their treatment in Darwin's original theory (Griesemer 2013, Wimsatt 1987). Whereas in Darwin (1859) they were inferred from observational generalisations, or, in the case of heritability, from an analogical argument, and were entangled with other variables, Lewontin extracted the three principles by isolating them and gave them an explicit and quite rigorous definition, so that they could be as independent as possible from the empirical context in which they originated. This operation increased the quasi-axiomaticity and the empirical shielding of these theoretical components, in comparison to Darwin's formulation, but also in comparison to the Modern Synthesis.

Lewontin's formalisation was aimed at showing that evolutionary change does not necessarily happen only at the level of individual organisms, as was implied by the formulation of evolution by natural

91 By generalising over this claim, it may be argued that every functional extension of a principle increases its generative potential, almost by definition, although this would require a separate argument. However, the contrary is clearly not valid. Not all increases in generative potential of theoretical or other kinds of epistemic components result from endogenization or functional extension nor, more in general, do they lead to cases of internalisation.

selection provided by the neo-Darwinian framework. His three principles abstract and idealise away, among others, from the assumption that natural selection only operates on *individual organisms*, understood as genetically discrete and homogeneous units. Variation, differential fitness, and heritability represent the enabling conditions for natural selection to operate on *populations of individuals*, where 'individual' does not refer to any privileged type of biological entity, but it can be applied at all the different hierarchical levels (molecules, cells, organisms, groups). Once we recognise that this chapter in the epistemic history of evolutionary biology is closely connected with the case of endogenization of individuality analysed in section 3.2.2, Lewontin's formalisation can be interpreted as the first stage of a process of endogenization. Its second stage was the functional extension of the principles to explain hierarchical organisation and, eventually, internalise individuality, that is, describe it as a result of natural selection and not any longer as a background presupposition.[92]

More precisely, Lewontin's formalisation crucially contributed to fixing the epistemic function of the three core Darwinian principles so as to make them the most relevant explanatory variables within the theory, thus increasing their quasi-axiomaticity, in comparison both to previous, less formal formulations, and to other variables, which were fixed as background presuppositions of natural selection (e.g., individuality, environment, etc.). On the contrary, the quasi-axiomaticity of these exogenous variables decreased, since they do not figure any longer as fixed background conditions necessary for natural selection to operate. At the same time, his formalised principles abstract away from the empirical details that were so important for Darwin to formulate his theory in the first place and, therefore, increased their empirical shielding.[93] The fixing of the epistemic function of the core Darwinian principles and their abstraction from specific empirical domains was required to achieve their functional extension as modelling principles for phenomena previously backgrounded. This, in turn, led to the internalisation of those phenomena, mirrored by an increase of the generative potential of the principles. The increase of their generative potential is an increase of the amount of conceptual dependencies that these principles *could* sustain (hence the notion of 'generative *potential*'), that is, of the amount of further conceptual tools depending on them that could be developed. This increase of generative potential is rooted in the phase of formalisation, since abstraction and idealisation made

[92] Cf. Okasha (2018), section 3.5 for more on the endogenization of hierarchical organisation.
[93] In the case of the principle of heritability, it may be pointed out that Galton already made available an abstract statistical concept of heritability and, therefore, that heritability had high empirical shielding a long time before Lewontin's formulation. However, the reason behind Galton's statistical concept of heritability was that of black-boxing the mechanism of genetic inheritance, that had not yet been (re)discovered. In my view, this epistemic move is quite different from that of formulating a general principle of heritability through abstraction from known details, i.e., detaching it from the specific case of biological reproduction and genetic inheritance, as in the case of Lewontin's formulation.

the principles more general and fixed and, thus, 'made room', so to speak, for a larger justificatory potential for further representational resources. Yet, the principles acquired a higher generative potential during the stage of their functional extension by which the internalisation of background variables was achieved. Yet, as I will show in the next section, this increase of generative potential has further effects in a subsequent phase.

In a nutshell, I am suggesting that Lewontin's formalisation directly underpinned the functional extension of the principles of variation, differential fitness, and heritability towards some empirical phenomena represented by variables that were formerly idealised away as background conditions (previously themselves characterised by a high quasi-axiomaticity). More specifically, Lewontin's formalisation, by increasing the quasi-axiomaticity and empirical shielding of the core Darwinian principles, contributed to increasing their generative potential and making them the 'endogenizers', thus enabling their functional extension for the purposes of modelling the formerly backgrounded variables of hierarchical organisation and individuality. At the same time, these variables – previously representing exogenous presuppositions of evolution by natural selection as formulated within the neo-Darwinian framework – became 'endogenized'. This means that their epistemic function changed from being fixed variables idealised away in the background of natural selection theory, to represent domains of phenomena describable by means of the endogenizers, i.e., explainable as products of natural selection. In sum, endogenization can be characterised as a simultaneous change of two components within a theoretical framework. On the one hand, the endogenizers' epistemic function is set fixed and extended to novel domains while, on the other, the endogenized components are transformed from background idealised variables into novel 'explananda' of the framework.

This claim of a connection between Lewontin's formalisation and the endogenization of hierarchical organisation and individuality may be generalised. I would like to suggest that the three principles identified by Lewontin are involved as endogenizers in other cases, in addition to those of hierarchical organisation and individuality. In the case of niche construction theory, where the variable of selective environment is endogenized since it is, at least in part, accounted for as a product of natural selection, we can see how the three Lewontinian principles were functionally extended. According to niche construction theorists, phenotypic differences in the modes of interaction between organisms and environment (variation), that are considered to be heritable (heritability), although not only through biological reproduction, provide a selective advantage (differential fitness) to the organism-environment niche. More specifically, the core Darwinian principles are endogenizers in that their theoretical scope is extended so as to describe a phenomenon that previously was not considered as influencing evolutionary dynamics, that is, niche-constructing activities. Since niche-constructing activities change the environment, and these changes produce further transformations both in the distribution of phenotypic variations and in the selective pressures cumulatively inherited by future generations, the variable of selective environment is the endogenized component, because rather than being assumed as a background assumption it is explained in terms of the endogenizers. Certainly, much more work is required to consolidate the claim of a relationship between the formalisation of these principles and endogenization as a process of theory extension, by looking at the details of each case and at the changes in the epistemic role of both the endogenizers and the endogenized. Yet, in the next subsection, I must move on to consider how the functional extension of theoretical principles and their increased generative potential impact their coordination with the level of phenomena.

3.3.3 Coordinating 'heritability' and 'inheritance' in niche construction theory

3.3.3.1 From functional towards semantic extension

As I mentioned above, the functional extension of the endogenizers corresponded to an increase of their generative potential, that is, to an increase of the amount of conceptual dependencies that they can support or, in other words, of the amount of downstream conceptual components that depend on them. Consequently, the functional extension of the endogenizers influenced certain conceptual components that partly depend on them and that have a fundamental role in the coordinating the endogenizers themselves with the phenomena that they are supposed to represent.

What I will argue in this subsection is that the functional extension of the core Darwinian principles had effects on the meaning of certain fundamental concepts involved in the coordination between theory and phenomena. More specifically, I will focus on how the functional extension of the principle of heritability, expressing an abstract intergenerational correlation in variation distributions, pushed for a semantic extension of the concept 'inheritance', which identifies the concrete mechanism(s) of intergenerational transmission and retention of variation. In other words, the scope extension of the theoretical role of the principle of heritability (together with the other two principles) to model and, eventually, internalise certain variables, required and extension of the referential scope of 'inheritance'. Yet, as I will show, while the principle may have constrained the extension of the meaning of 'inheritance', it did not delimit its precise scope. At the same time, current discussions on the correct meaning of 'inheritance' reflect a condition by which a unanimous semantic extension of the term has not yet been achieved. From these elements, I conclude that the increase of the generative potential of the principles, mirroring their functional extension, partly explains the success of endogenization as a theoretical strategy (in combination with epistemic values such as simplicity and scope). However, it is only by (re-)coordinating the principles with the endogenized phenomena via certain fundamental concepts, that is, by achieving a *justified* semantic extension of these concepts, that endogenization is fully accomplished as internalisation of phenomena. To illustrate these aspects concretely, I now set out to discuss the functional extension of the principle of heritability and its effects on the concept of 'inheritance' with respect to the endogenization of selective environment in niche construction theory.

As I mentioned at the end of section 3.3.2, in niche construction theory the core Darwinian principles can be regarded as endogenizers of selective environment, which is the endogenized variable. Yet, the resulting endogenization of selective environment seems to rely on a different understanding of 'inheritance' than the one of genetic inheritance assumed by the neo-Darwinian framework, i.e., the transmission of genetic material from parents to offspring during reproduction, taken to be responsible for the phenotypic heritability (i.e., more similarity between parents and descendants than between these and other members of the population). In fact, niche construction theorists assume another dimension of inheritance, on top of the genetic one: "Selected habitats, modified habitats, and modified sources of natural selection in those habitats are also transmitted by the same organisms to their descendants, as a consequence of their niche-constructing activities, through a second general inheritance system in evolution, ecological inheritance" (Odling-Smee 2007: 278). Ecological inheritance is an 'indirect' type of inheritance, in contrast to the direct transmission of genes from parents to offspring, since it does not involve directly a generational event, but happens through the medium of the environment. Why then calling it inheritance? What is the justification for extending the referential scope of this term? If we look at the debates on the meaning of 'inheritance', we quickly acknowledge that this is far from being a settled issue in evolutionary biology. In what follows, I briefly rehearse the recent history of this concept until contemporary controversies, to show how different interpretations of the principle of heritability are held fixed in order to provide alternative characterisations of the notion of inheritance.

3.3.3.2 Conceptual controversies over 'extended inheritance'

As I mentioned above, Lewontin's principle of heritability merely implies a statistical correlation in the retention of changes of phenotypic variation across generations.[94] What Lewontin had in mind was an abstract *descriptive* characterisation of a correlation in terms of the statistical measure that biologists use to model evolutionary dynamics, rather than a *causal* notion of inheritance as a set of

[94] Here, I am referring to his 1985 formulation of the principle, in which he corrected his 1970 version where he expressed the principle of heritability in terms of 'heritability of fitness'. Cf. above, footnote 88.

physical systems or processes responsible for that state. According to the neo-Darwinian framework, one single mechanism 'realises' this statistical correlation: genetic inheritance, i.e., the transmission of stretches of DNA between parents and offspring via reproduction. In this sense, it may be said that, within the neo-Darwinian framework, the term 'inheritance' only refers to the mechanism of genetic inheritance and, in this way, it enables the representation of physical states (i.e., actual intergenerational transmission of phenotypic variation) in terms of a statistical correlation (i.e., the intergenerational correlation of variation distribution expressed by the principle of heritability) via a single physical process (genetic inheritance). In other words, the identification of inheritance with the causal mechanism of genetic inheritance is a central piece that, within the neo-Darwinian framework, enables the univocal coordination between physical states and their abstract representation as a statistical correlation.

However, this identification started being challenged by the rise of new subdisciplines of evolutionary biology, such as developmental systems theory, gene-culture co-evolution theories, and epigenetics. Scientists working in these fields refer by the label of 'non-genetic inheritance' to other causal factors (variously conceptualised as 'systems', 'channels', 'mechanisms', or 'processes') that influence the retention of changes in variation distributions within and across generations, such as epigenetic, behavioural, ecological, and cultural factors. Consequently, many proposals for an extended concept of inheritance have been put forward (Danchin et al. 2011, Griffiths & Gray 1994, Jablonka & Lamb 2005, Odling-Smee 2007, Sterelny 1996). However, the proliferation of putative inheritance systems as a reaction to the restrictive neo-Darwinian identification of inheritance with the genetic mechanism, raised the worry of obtaining, as a result, an over-inclusive concept of inheritance. This worry refers to two possible deleterious effects. In the first place, a notion of inheritance that extends to encompass too many phenomena might create problems when it comes to formal modelling of evolutionary dynamics. The neo-Darwinian identification of genetic inheritance as the only mechanism preserving the retention of changes in variation distributions among generations is – now we know it – a strong idealisation. Yet, this idealisation served well the modelling purpose, as testified by the great achievements of population genetics in the 20th century (Odling-Smee 2007). The second issue comes from the realisation that most of the proposals to extend 'inheritance' assume that every form of transmission is a form of inheritance (Merlin 2017). In other words, these reconceptualisations have abstracted away from the identification of gene transmission via reproduction as the only inheritance mechanism, since this identification was empirically inadequate, but this abstraction may lead them astray to be over-inclusive.

Reconceptualising 'inheritance', in my view, is a necessary phase after the functional extension of the principle of heritability. This extension was aimed at modelling phenotypic similarity that was previously idealised away in the form of background variables. In the case of niche construction theory, it aimed at capturing changes in the distribution of phenotypic traits that result from niche constructing activities. Yet, the functional extension of the principle of heritability stretched and eventually compromised the coordination between heritability, expressed as an intergenerational statistical correlation in variation distributions, and the actual presence of more similarity between certain members of a population than among all other members, i.e., the phenomenon that the principle of heritability describes in abstract statistical terms. More specifically, the functional extension of the principle of heritability 'broke' the semantic identification of 'inheritance' with genetic inheritance, which was the crucial piece of the coordination, and required a semantic extension of the term, so as to include more processes responsible for phenotypic similarity. In this sense, the functional extension of the principle of heritability pushed for a reconceptualization, or semantic extension, of 'inheritance', albeit without providing clear boundaries for its extension (fig. 3.2a and 3.2b).

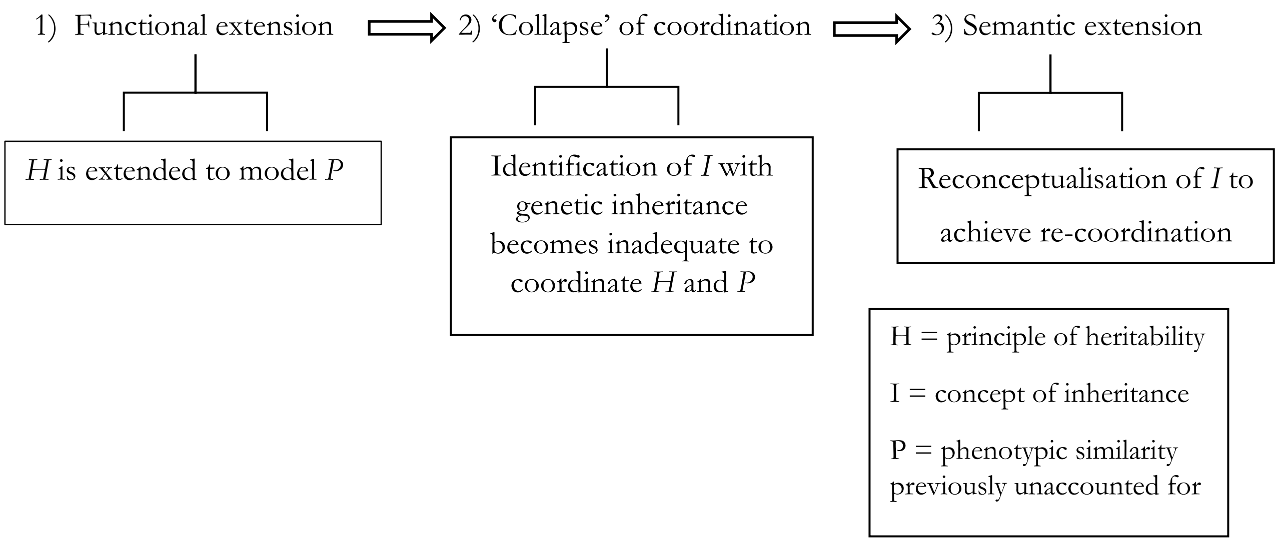


fig. 3.2a: Unfolding of the effects of the functional extension of heritability on the concept of inheritance. The arrows indicate the succession of phases

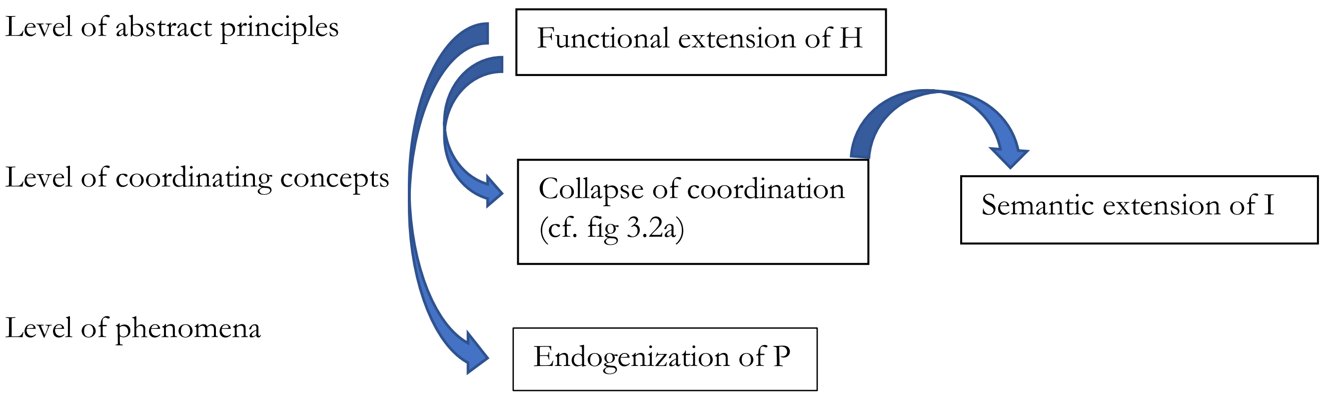


fig. 3.2b: representation of the effects of functional extension with respect to the level of theoretical abstraction of the epistemic components and of their relationships

According to Merlin (2017), current proposals for extending 'inheritance' fail to do this properly, since they end up including types of transmission that are not in line with what she calls the 'theoretical role of inheritance' in natural selection theory. For instance, forms of horizontal and oblique transmission, such as epigenetic lateral gene transfer or cultural inheritance, which bypass inter-generational transmission via reproduction, are merely ways of acquiring new variations, and not of cumulating already existing variations and preserving their continuity across generations, as the theoretical role of inheritance would prescribe. Merlin's point is worth dwelling on. The problem she raises is how to conceptualise 'inheritance' in the light of empirical evidence of new causally relevant factors, other than genetic ones, influencing the retention in changes of variation distributions across generations. This conceptualisation should include all those physical realisers, understood as mechanisms, processes, systems, or channels, which contribute to inheritance, but leave out those physical realisers that merely contribute to the acquisition of variation. To do that, she appeals to the theoretical role of inheritance with explicit reference to Lewontin's formulation of the principle of heritability as one of the enabling conditions for natural selection. Lewontin's formulation of this principle states that a statistical correlation holds between parents and offspring and, therefore, according to Merlin, it requires that transmission must occur in a parent-offspring lineage, and not, for instance, within the same generation. In addition, these correlations relate to the contribution of parents and offspring to future generations, therefore implying continuity across generations and the cumulation of variation across time. In Merlin's view, if we assume that a mechanism of transmission, without preserving continuity and the cumulation of variation, is sufficient for inheritance, we make the mistake of including physical mechanisms or processes which only realise acquisition of variation rather than inheritance. Merlin's conclusion is that the notion of inheritance must be extended beyond mere genetic inheritance, but not so much as to include mechanisms that do not involve the transmission of some generational material through generational events. Inheritance, she argues, requires reproduction as a mechanism of transmission, since it is the only mechanism that preserves continuity and cumulation of changes in variation distributions. Transmission mechanisms without reproduction do not serve these purposes, therefore, cannot be labelled as 'inheritance systems', given the theoretical role of 'inheritance'. Rather, they relate to phenomena better described in terms of acquisition of new variation.

It is interesting to point out how Merlin's argument is based on holding fixed what she calls a 'theoretical role of inheritance' to then mould the extended meaning of 'inheritance' so as to include some transmission mechanisms and exclude others. Her argument seems to presuppose a specific reading of the concept 'inheritance' out of Lewontin's abstract formulation of the principle of heritability. From heritability as a statistical description of intergenerational changes in variation distribution, she derives a 'theoretical role of inheritance' – based on the necessary presence of a parent-offspring lineage, of continuity across generations, and of the cumulation of variation across time – as a guiding principle to extend the notion of inheritance. However, the reading of what counts as a causally relevant factor for intergenerational changes in variation distribution can hardly be derived from an abstract statistical description. In order to tease out indications of causal relevance from what she calls the 'theoretical role of inheritance' derived from Lewontin's principle of heritability, Merlin must already presuppose some physical interpretation of that abstract description. Yet, the requirements that make up Merlin's 'theoretical role of inheritance' do not seem to be uncontroversial, and different interpretations of the principle of heritability seem to guide different choices concerning the extension of the term 'inheritance'. For example, Charbonneau (2014) argues that parent-offspring lineages are not necessary for a population to undergo Darwinian natural selection, as Lewontin's principle of heritability would seem to require, according to Merlin. Therefore, in Charbonneau's view, 'inheritance' should not refer only to vertical transmission, i.e., transmission between parents and offspring, as the only system which preserves transgenerational retention of changes in variation distributions. Rather, generation and memory, which Charbonneau (2014: 739) defines "without reference to a specific biological ontology of entities and processes", suffice to guarantee transmission, continuity, and cumulation of changes in variation distributions. Generation and memory can be physically realised by mechanisms other than genetic transmission or transmission via reproduction, as in the case of diffused ecological inheritance. Although Charbonneau emphasises the contrast between his view and Lewontin's formulation of the principle of heritability, the former could rather be considered as its development, since it pushes even further Lewontin's formalisation by explicitly abstracting away from parent-offspring lineages, while it preserves the functional role of heritability as abstract statistical characterisation without conflating it with the causal meaning of 'inheritance'.

3.3.3.3 Back to heritability and inheritance in niche construction theory

Let us now go back to niche construction theory and the endogenization of selective environment. As I mentioned above, niche construction theory extends the notion of inheritance to encompass ecological inheritance in addition to genetic inheritance. Ecological inheritance is an indirect system of inheritance, since it does not require vertical transmission nor reproduction. Rather, it happens through the medium of the environment, which is modified by the niche-constructing activities of the organisms. The cumulative effect of these modifications influences future generations in that they directly produce durable changes in the distribution of certain phenotypic variations or, more deeply, by modifying selective pressures.[95] Yet, one might wonder – on the lines of Merlin's argument – whether that counts as inheritance, and not as acquisition of new variations. According to Odling-Smee (2007: 280), ecological inheritance is indeed inheritance "because some of the environmental consequences caused by the repeated niche-constructing activities of multiple generations of organisms in their environments *accumulate* or *persist* in environments across generations".[96] The function of cumulation and continuity in the transmission of changes in variation distributions is satisfied by ecological inheritance, even though the transmission has high fidelity only in its interaction with genetic inheritance (Odling-Smee & Laland 2011). This quote illustrates that niche construction theorists have reconceptualized the notion of inheritance by extending it in such a way that it refers also to the system of ecological inheritance. At the same time, their semantic extension of 'inheritance' is guided by what they assume is its core meaning, that is, as the concept enabling the identification of those systems that contribute to the retention of changes in variation distributions, by means of transmission, cumulation, and continuity.

Again, their reconceptualization of 'inheritance' can be viewed in its connection with their interpretation of the principle of heritability. Niche construction theorists functionally extended the core Darwinian principles to use them to model niche-constructing activities in terms of evolutionary dynamics, thus leading to the endogenization of the variable of selective environment. This functional

[95] Charbonneau (2014) focuses on the case of diffused ecological inheritance as a mechanism which influences trait variation, rather than selective pressures, whereas Odling-Smee (2007) and Odling-Smee & Laland (2011) discuss niche inheritance as the interaction between genetic and ecological inheritance, where the latter is understood mainly as inheritance of selective pressures. These two understandings of ecological inheritance are not in contrast but, rather, complementary (Charbonneau, personal communication). For the purposes of my discussion, it is not necessary to deal with the intricacies of this distinction.
[96] Emphasis mine.

extension precedes the achievement of a justification for the new extended meaning of 'inheritance' as inclusive of ecological inheritance. Yet, the functional extension had an impact on the coordination between the core Darwinian principles and the phenomena endogenized. As I mentioned above, the extension of the principle of heritability to describe (kinds of) phenotypic similarity that could not be causally explained in terms of genetic inheritance (by niche construction theory and other branches of evolutionary biology), highjacked the identification of 'inheritance' with the mechanism of genetic inheritance. This identification was, within the neo-Darwinian framework, crucial to the coordination of heritability as a statistical correlation with the concrete patterns of phenotypic similarity it is supposed to represent, since it posited one causal mechanism by which those patterns are realised. The 'breakdown' of this identification required the search for a new, more abstract and inclusive concept of inheritance that would mirror the functional extension of the principle of heritability to represent more (kinds of) phenotypic variation than before. Hence, the conceptual controversies on 'extended inheritance' that I described above. Within this context, also niche construction theorists, in the same way as Merlin and Charbonneau, based their extended notion of inheritance on their interpretation of the principle of heritability. More specifically, they based it on their need to functionally extend that principle to model certain phenomena, in this case, the transmission and retention of traits via niche-construction activities (i.e., a kind of phenotypic similarity that was excluded as such by previous theorising). Therefore, their semantic extension of 'inheritance' reflects their modelling practice and their endogenization of selective environment. At the same time, it is the result of their attempt to re-coordinate the principle of heritability with the extended domain of phenomena that they model in evolutionary terms by reconstituting a concept that crucially contributes to that coordination. The epistemic function of the principle of heritability is, therefore, a starting point to re-coordinate 'inheritance' with the extended domain of phenomena, that is, to identify which mechanisms or systems count as inheritance mechanisms or systems, and which do not. This epistemological move is not peculiar to the way niche construction theorists provided their formulation of extended inheritance but is reflected by the larger debate on the extended concept of inheritance, both in permissive and in restrictive views, as I discussed above.

Controversies over the adequate semantic extension of the term 'inheritance' are still hotly debated. I have suggested that these controversies result from the impact of the functional extension of the principle of heritability, together with those of variation and differential fitness, aimed at internalising phenomena previously idealised away as background variables. The functional extension of the principle of heritability compromised the identification of the concept of inheritance with the mechanism of genetic inheritance. This identification was crucial to the coordination between the abstract principle of heritability and the concrete phenotypic similarity it was supposed to describe in the neo-Darwinian framework, since it justified the prescription that all phenotypic similarity is due to genetic inheritance. Extending the theoretical function of heritability led to the requirement of extending the meaning of 'inheritance', in such a way that further inheritance mechanisms would be included. In this way, a novel form of coordination between the principle of heritability and the endogenized phenomena would be achieved, i.e., by means of a univocally extended notion of inheritance. In this sense, however, the notion of inheritance still lacks a unanimous semantic extension. As I showed above, different interpretations of the principle of heritability prescribe different ways of working out the referential scope of 'inheritance' by coordinating it with the physical transmission mechanisms that best fit both evidence and theory. In the case of niche construction theory, the referential scope of 'inheritance' has been stretched to include the mechanism of ecological inheritance in the light of the novel modelling purposes pursued by the discipline and justified by the functional extension of the core Darwinian principles to new empirical domains. In current discussions over an extended evolutionary synthesis, different attitudes towards restricting or expanding the referential scope of 'inheritance' seem to reflect different modelling and explanatory purposes and, thus, may be subject to a pluralistic interpretation.

3.3.4 Constitutivity increases with endogenization

Up to this point, I have suggested that endogenization as a strategy of theory extension in evolutionary biology comprises different stages. In the first stage, formalisation increases the quasi-axiomaticity and empirical shielding of the endogenizers, (principles of variation, differential fitness, and heritability). Thanks to their formalisation, the theoretical function of the principles can be extended so as to model phenomena that were previously idealised away as background variables of evolution by natural selection. This functional extension is mirrored by an increase of the generative potential of the principles, that is, of their capacity of sustaining conceptual dependencies. As a result of this second stage of functional extension, some phenomena previously idealised away in the form of background variables were internalised, that is, explained away by the theory of natural selection. However, the extension of the principles had an impact on certain concepts that were crucial to the coordination between theory and phenomena, i.e., that enabled the abstract principles to represent real phenomena within the theory of natural selection. More specifically, I showed that, in the case of the principle of heritability, controversies on the semantic extension of the term 'inheritance' are still open, while the core Darwinian principle of heritability has long been used to model phenotypic variation that was not conceived as heritable within the neo-Darwinian framework. In fig. 3.3, I extend the table represented in fig. 3.1 to include what I consider the third, final stage of endogenization, i.e., the semantic extension of certain concepts that are crucial to the (re-)coordination between abstract principles and phenomena.

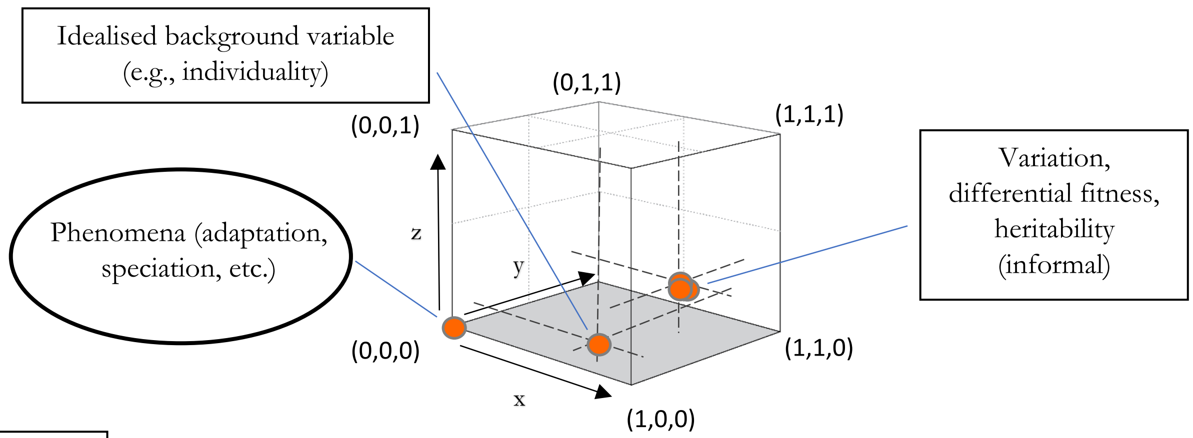
Stage	1) Formalisation	2) Functional extension	3) Semantic extension
Outcome	Abstraction of the principles → ⇧ generality Idealisation → fixing of their epistemic function	Use of the principles to model previously backgrounded phenomena → - internalisation of phenomena - collapse of coordination	Extension of referential scope of coordinating concepts → - justification of (re-)coordination - successful internalisation
Parameters	⇧ empirical shielding (of principles) ⇧ quasi-axiomaticity (of principles)	⇧ generative potential (of principles)	Possibly ⇧ quasi-axiomaticity, empirical shielding & generative potential (of coordinating concepts)

fig. 3.3: Table summarising the outcomes of the process of endogenization from formalisation to semantic extension, via functional extension, and their effects on the empirical shielding, quasi-axiomaticity, and generative potential of the epistemic resources involved. Whereas formalisation and functional extension regard the three core Darwinian principles, semantic extension refers to core conceptual resources involved in the coordination between the principles and phenomena (e.g., the concept of inheritance)

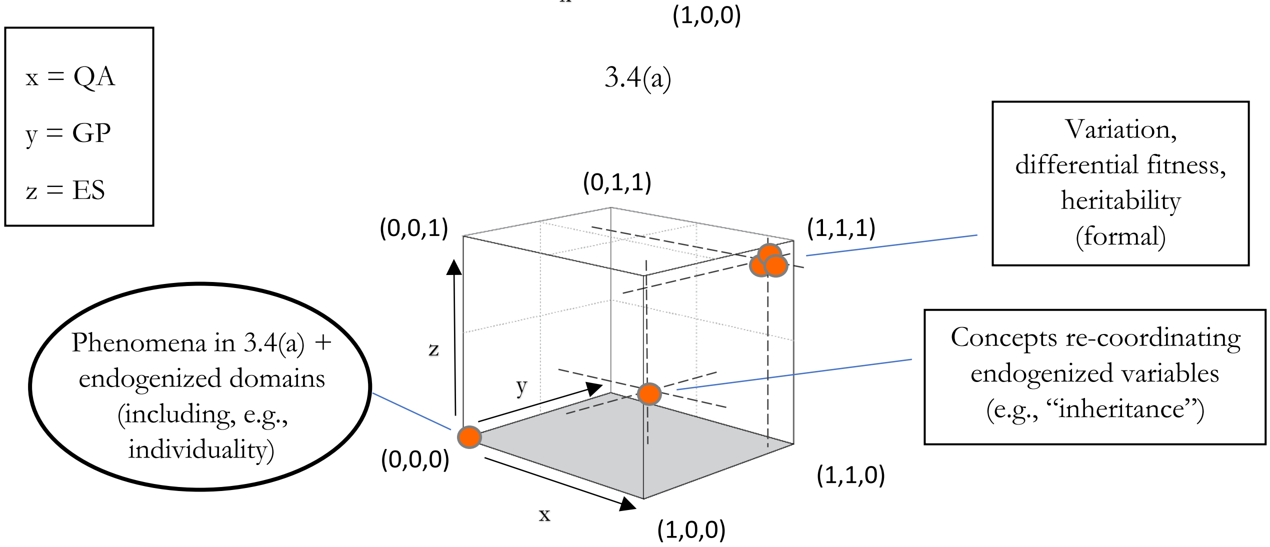
Prima facie, the core Darwinian principles seem to have a different epistemic function than those characterised by Friedman as constitutive principles, as I discussed them in chapter 2. For instance, Newton's laws of motion provided the tools for formulating and testing the law of gravitation, in that they conceptually coordinated the mathematics of the calculus with the concrete phenomena that the law of gravitation captures. On the other hand, the principles of variation, differential fitness, and heritability provided an explanatory structure that, once abstracted away from empirical details, could be applied to further phenomena, even to those that were backgrounded in the first place. In this sense, endogenization itself represents a crucial difference between the core Darwinian principles and Friedman's examples of constitutive principles. The generative potential of Newton's laws enabled the formulation of a unified theoretical representation of a set of phenomena by means of the law of gravitation, whereas, in the case of the core Darwinian principles, their increased generative potential mirrored a progressive expansion of the theoretical role of those principles to further domains.

Still, the core Darwinian principles express relatively abstract theoretical conditions that, at different stages of development in evolutionary biology, constrained certain conceptual possibilities. On the one hand, once explicitly formalised, their functional extension was required to internalise phenomena previously idealised away as background variables; on the other, they were held fixed as enabling conditions for evolution by natural selection, although with different interpretations, so that the extension in meaning of certain core concepts used to coordinate them with phenomena could be discussed, as I showed in the case of 'inheritance'. In other words, the endogenizers constrain from top-down, that is, from the theoretical level, the process of (re-)coordinating certain core concepts with the extended domain of phenomena to which the scope of natural selection theorising has been extended. This means that semantic extension can be viewed as a final stage of endogenization, one in which scientists support the modelling goals that motivated the functional extension of the principles in the first place with a more extended discussion of certain conceptual issues, which come about as effects of functional extension on the coordination between theory and phenomena. In this sense, the principles provide some crucial theoretical fixed points to investigate the limits of the semantic extension of the coordinating concepts, as I suggested in the case of 'inheritance'. Across the three stages of the process of endogenization – formalisation, functional extension, and semantic extension – the empirical shielding, quasi-axiomaticity, and generative potential of the core Darwinian principles has increased, thus increasing their constitutivity.

Fig. 3.4(a) represents the relationship between the domain of phenomena accounted for by natural selection in the theoretical framework developed by Darwin (1859) and the principles of variation, heritability, and differential fitness, which were here informally formulated. Given the phenomena accounted for by Darwin in terms of evolution by natural selection (framed as the 0,0,0 point), the three principles have a low empirical shielding because they directly stem from the large amount of empirical observations on which Darwin generalised. Their quasi-axiomaticity is relatively high, because their function as explanatory resources is fixed within the framework, even though not maximal, since many other variables appear, with which they interact at the same level of generality (e.g., struggle for survival, sexual selection, etc.). Their generative potential is also relatively high since, within the framework, they sustain a large amount of conceptual dependencies. In contrast, figure 3.4(b) represents the relationship between the domain of phenomena explained by natural selection in the post-Modern Synthesis and post-Lewontin formalisation evolutionary biology, in which strategies of endogenization are pursued and the three principles are now more formally expressed. In this representation, the geometrical distance between the domains of phenomena conceptually framed as subject to natural selection – that includes those in figure 3.4(a) plus those of endogenized domains – increases due to two factors. On the one hand, the empirical shielding and quasi-axiomaticity of the principles have increased, given the more abstract and idealised formulation of the principles. On the other, the generative potential of the principles has also increased, not only due to the larger number of phenomena accounted for by natural selection theory, as a result of endogenization, but also because they function as fixed theoretical points of reference for the clarification of core concepts, such as that of inheritance.



3.4(a)



3.4(b)

This is only a very schematic representation of the differences in the degree of quasi-axiomaticity, empirical shielding, and generative potential of the core Darwinian principles, i.e., the endogenizers, at two different stages of the history of evolution by natural selection as a theory. My interest here is to show how the constitutive character of certain components can vary in the epistemological macro-dynamics, depending on the use of certain theoretical strategies such as endogenization. Nonetheless, my conclusions are constrained by the fact that the figures are highly abstract and my analysis is limited to one case only, that of 'inheritance' as one of those conceptual components that enable the re-coordination of the principles with the novel empirical domains of application, but whose meaning is constrained by the functional extension of the endogenizers.

Yet, the case of inheritance is crucial to my suggestion that the core Darwinian principles are not just more general, but more constitutive. Whereas Okasha (2018) argues that endogenization mirrors a more general character of the core principles of natural selection, compared to their role in the Darwinian and neo-Darwinian framework, I suggest that they become more constitutive. I share Okasha's view that the more abstract nature of these principles does not overturn the importance of contextual information in the new (formerly backgrounded) empirical domains. In order to apply these principles to conceptually frame novel phenomena, scientist must resort to a great deal of further contextual knowledge, so that the abstract theoretical resources can have meaning in the new empirical context. It is in this sense – as I said – that the core theoretical principles must be re-coordinated with the empirical phenomena. However, my account elucidates two aspects that Okasha's interpretation of endogenization does not. On the one hand, the process of endogenization increases the quasi-axiomaticity, empirical shielding, and generative potential of the endogenizers, while also influencing the endogenized. The formerly idealised background variables switch from having a high quasi-axiomaticity to a low one, because they are explained away in terms of the endogenizers, as represented in the pictures 3.4(a) and 3.4(b) with respect to individuality. On the other hand, the core Darwinian principles are constitutive, and not just general, because they are held as fixed theoretical points of reference that constrain the attempts at reconceptualising some core concepts of the theory that partly depend conceptually on the principles themselves. As I showed for the case of inheritance, these concepts are themselves crucial to the coordination of the principles with the level of phenomena and have relative degrees of quasi-axiomaticity, empirical shielding, and generative potential, which may further increase with their semantic extension.

CHAPTER 4

4.1 Measurement, circularity, and the meaning of quantity concepts

4.1.1 Understanding coordination in current epistemology of measurement

So far, I have discussed examples of coordination as a process or an achieved condition that involves the use of some theoretical principles to justify a certain referential relationship. In one case, i.e., that of Newton's laws of motion, these principles are involved in coordination because they justify the referential relationship between a mathematical representation of an empirical regularity and the concrete phenomena that it represents. In the second case, i.e., that involving heritability and inheritance in the endogenization of selective environment, the semantic extension of 'inheritance' is justified in the light of an already extended epistemic function of one or more theoretical principles. These principles fulfil their epistemic function by exhibiting high degrees of quasi-axiomaticity, empirical shielding, and generative potential, and, therefore, are to be deemed as constitutive to a significant extent.

However, as I explained in section 2.1, the term 'coordination', as it is found in classic and contemporary literature, is used with respect to distinct, albeit closely related, issues. The meaning of 'coordination' that I have assumed so far refers to the general issue of providing the conditions of applicability of mathematically expressed laws, principles, or theoretical concepts to concrete phenomena. Another – narrower – meaning of 'coordination' has been revived by recent literature in epistemology of measurement, with reference to a specific metrological concern, as I already mentioned when introducing van Fraassen's view in chapter 1.[97] This concern is the problem of how scientists justify their belief that certain measurement procedures identify the quantity of interest in the absence of independent methods to assess them. Contemporary epistemologists of measurement refer to this issue as to the 'problem of nomic measurement' or the issue of 'coordination' between the meaning of quantity concepts and measurement procedures.[98] In order to solve this problem, Chang (2004) and van Fraassen (2008) suggest considering the meaning of quantity concepts as emerging along the process through which a form of coordination with the measurement procedures

[97] Cf. above, section 1.2.3.2.

[98] Chang (2004) labels this issue the 'problem of nomic measurement', followed by other scholars (e.g., Boumans 2005, Cartwright & Bradburn 2011, Sherry 2011). However, in the SEP page on 'Measurement in science', Tal (2017a: section 8.1) highlights that "this circularity has been variously called the "problem of coordination" (van Fraassen 2008: Ch. 5) and the "problem of nomic measurement" (Chang 2004: Ch. 2)". This statement mirrors the rather interchangeable use of both expressions in the literature. Since my goal in this chapter is not one of conceptual cleaning, I will not discuss the legitimacy of the use of these expressions.

is achieved, thus emphasising the constructive aspect of the dynamics between theorising and measuring. Their work stimulated the production of a wealth of case studies focusing on this issue across scientific disciplines (Barwich & Chang 2015, Germain 2014, McClimans 2013, McClimans et al. 2017, Michel 2019, Ruthenberg & Chang 2017, Tal 2011). Further contributions have significantly advanced the study of coordination with respect to measurement. These works provide in-depth analyses of epistemic components internal to the measurement process, e.g., calibration, and clarify crucial epistemological distinctions such as the one between instrument readings and measurement outcomes (Boumans 2007; Frigerio et al. 2010; Giordani & Mari 2012; Mari 2003; Tal 2013, 2016a, 2016b, 2017b, 2019).

In this chapter, I focus mainly on the specific coordination between measurement procedures and the meaning of quantity terms. On the one hand, I assume that this narrower sort of coordination is a special case of the general kind of coordination between abstract representational resources and concrete phenomena. The issue of coordination between quantities, their concepts, and measurement procedures may, in fact, be viewed as a special case of the more general issue of providing the conditions of applicability of theoretical terms to concrete phenomena. Yet, the distinction between these two scopes of coordination is useful because it helps emphasising the specific contribution of the narrower kind of coordination to the achievement of the more general one, as it will be clear from the discussion of the historical case that I analyse in this chapter. On the other hand, focusing on the narrower case of coordination will allow me to suggest that non-theoretical epistemic components can also be constitutive. More specifically, the analysis of measurement practices and of their material aspects will lead me to suggest that scientific instruments can also be constitutive, to some degree, within the restricted scope of the coordination between measurement procedures and the meaning of quantity terms.

In the rest of this section, I will firstly introduce the issue of coordination in epistemology of measurement. To do that, I will briefly present Chang's (2004) approach, since he and van Fraassen (2008) developed the two most influential accounts tackling this issue in the recent literature. Yet, for my purposes, most of their differences will be negligible. Secondly, I will discuss the conceptual distinction between two kinds or scales of coordination. Finally, I will present, in a systematic way, some categories recently discussed by epistemologists of measurement, which enable a rich and fruitful analysis of the internal components of measurement. These categories will be useful to reconstruct, in section 4.2, Ohm's experimental practice. The case of Ohm's research on electrical conductivity is particularly suitable to illustrate the process through which the narrower kind of coordination is achieved. I reconstruct how this process unfolded and how its results were related to the more general kind of coordination, i.e., the one between Ohm's law and the empirical regularity that it represents. Finally, in section 4.3, I build on the idea that one scientific instrument involved in Ohm's measurement practice had, together with other epistemic components, a constitutive character, and I develop systematically the notion of a 'material' kind of constitutivity, so that it can be expressed in terms of my account.

4.1.2 The problem of circularity in measurement and Chang's solution

Measuring practices have a pervasive role in establishing evidence and providing certain specific conditions for the applicability of abstract representations to phenomena. To measure some quantity, we often infer its value from the values of other quantities, as when we 'read' the temperature of a room out of the length of the mercury column in a thermometer hanging on the wall. This inference is based on knowledge of the physical law that describes the relationship between the quantities of temperature and length in a specific physical interaction. Such a *measurement law* represents the measurement interaction through which the desired quantity is inferred from other quantities (Chang 1995). But how do we get to know the form of the function that relates values of temperature and length? This question is the source of a conundrum: we need to *know* the law to assess the soundness of a measurement procedure. However, it is usually through measurements that we can establish and test empirical regularities. Clearly, if a certain measurement procedure has been used to establish the measurement law, its precision and accuracy cannot be assessed by means of the very same measurement law! Nonetheless, those cases are not infrequent in the history of science, in which it seems that "an understanding of measurement and what is measured is presupposed rather than established in our effort to assign meaningful values to the items in a scale" (McClimans 2013: 530). Put into a question, how can we be justified to believe that the measurement procedures we are deploying do measure the parameter of interest, if we lack independent ways to assess these procedures? Solving this problem of circularity equates to finding appropriate and independent sources of justification for a measurement procedure that assigns certain values to a quantity.

Chang (2004) focuses on the history of thermometry to develop his view of scientific progress and treats the problem of circularity I just described as a central issue related to it. According to Chang, the meaning of 'temperature' emerged from a process of cyclic feedback between theoretical advances and improvements in the standards for measuring it. At various stages of scientific progress, the same term 'temperature' was used with reference to different standards of observability and measurement, and it could assume different sets of values depending on the form of the relative measurement scale.[99] From his historical narrative Chang concludes that scientific progress follows cycles of epistemic iteration, that is, "a process in which successive stages of knowledge, each building on the preceding one, are created in order to enhance the achievement of epistemic goals" (Chang 2004: 45). Chang emphasises the historical character of this enterprise, based on the mutual refinement between theoretical background and measurement procedures. In other words, improvements at the level of measuring techniques contribute, on the one hand, to a better operationalisation of a quantity term which, in turn, enables the gathering of experimental data to support further theorising. On the other, theoretical advances allow for the refinement of the definition of the quantity term, by providing a guide to experimental measuring practices. This process embodies a sort of virtuous circle guided by the search for coherence between the various methodological and theoretical assumptions. According to Chang, the improvement of measurement standards happens very often quite independently from changes in theory, at least until theoretical advances are such that they can subsume and justify measurement procedures.

4.1.3 Two meanings of 'coordination'

One upshot of Chang's view is that a quantity cannot be assigned any value independently of both a measurement procedure that identifies its possible values and a form of coordination of the procedure with the rest of the conceptual apparatus in which the quantity concept is embedded. The solution to the problem of circularity in measurement – or 'problem of nomic measurement', in his terminology – boils down to progressively refining the provisional forms of justification of the measurement procedure developed in coherence with the theoretical knowledge and experimental techniques available at a certain stage of scientific development. Advancements in precision and reliability of measurement procedures occur via a process of mutual refinement with theoretical ones. Even among scientists, the consensus seems to lie on this historical and coherentist understanding of how coordination between quantity concepts and their standards of measurement is achieved (Johansson 2014). It is in this sense that both Chang and van Fraassen (2008) claim that the meaning of quantity terms is never a 'given', but it rather emerges along the various stages of this process.[100]

[99] Here, I refer to Stevens' (1946) standard fourfold classification of measurement scales: nominal, ordinal, interval, ratio. Cf. Mach (1896/1986) for a classic view on coordination and measurement with respect to temperature.

[100] Importantly, van Fraassen does not share Chang's emphasis on the extent to which measurement procedures can develop independently from theories.

As it is evident from my analysis in previous chapters, this sense of 'coordination', as the achievement of a stable justification for the identification of a quantity by means of a procedure for measuring it, does not exhaust the meaning of this notion. Solving the issue of circularity in measurement is only one way in which a form of coordination between abstract theoretical concepts and concrete phenomena can be established. Quite trivially, some theoretical concepts do not refer to measurable quantities. Most importantly, as we have seen, justification for the referential relationship between theory and phenomena can be obtained through epistemic operations that do not involve measurement, e.g., by means of theoretical justification, by relying on certain general assumptions about reality, or by any combination of the above. For this reason, I would like to spell out a conceptual distinction between two meanings of 'coordination'. This distinction reflects two uses of this notion that are different, yet tightly connected. As I will show with the Ohm case study, this distinction can be helpful for the analysis of specific scientific inquiries in which both kinds, or better, scopes of coordination are entangled. From now on, by *measurement coordination* I will indicate both the process by which a solution to the issue of circularity in measurement is achieved and the condition, resulting from this process, by which certain measurement standards reliably contribute to the identification of a quantity within a certain scientific framework. By *general coordination*, instead, I will keep referring to the broader process of coordinating abstract theoretical representations with concrete phenomena, abstracting away from the specific issue of circularity in measurement, and to the condition, resulting from this process, by which an abstract law, principle, or concept can justifiably represent an empirical phenomenon.

Achieving an improved form of measurement coordination involves several epistemic components (instruments, methodological, theoretical, and metaphysical assumptions, etc.). The activity of measuring is a complex one and it is entangled with experimentation, model-building, and theorising. For this reason, telling apart certain details of a measurement procedure can prove essential to understanding how measurement coordination is obtained and how it relates to general coordination. By 'details', here, I mean the different epistemic components of a measurement procedure that take on a peculiar justificatory role within a certain form of measurement coordination. These components can be specific theoretical, pragmatic, or metaphysical assumptions involved in the modelling of the measurement procedure itself, parts of the material experimental apparatus, further background theoretical commitments, etc. Indeed, Chang's work on temperature has provided a paradigmatic term of comparison with respect to its awareness of the importance of details to understand how scientists justify conceptual and metrological advances. Equally importantly, van Fraassen (2008) emphasises the central role of an epistemological distinction – the one between instrument readings and measurement outcomes – for the analysis of measurement coordination. Yet, recent literature has introduced an array of analytic concepts that facilitate the task of identifying the epistemic components that contribute to the justification for a certain form of coordination between quantities and the procedures to measure them.

In the next section, I introduce some metrological categories recently discussed by the flourishing literature in epistemology of measurement. Many of these works aim at clarifying the role and the relationships of the epistemic components interacting in the process that leads to measurement coordination. They can be viewed as deepening specific aspects that van Fraassen's and Chang's accounts, which I have taken as my starting point, dealt with to a lesser extent, given their broader scope. The conceptual distinctions that I am going to introduce will be crucial to my analysis of Ohm's scientific inquiry in section 4.2.

4.1.4 Disentangling measurement coordination

In fig. 4.1, I connect and integrate some recent conceptual developments in the literature on the main epistemic components typically involved in the process of measurement coordination. This graph is not supposed to be exhaustive but merely to provide a heuristic model for the study of measurement coordination. After the picture, I provide an explanation of the components and of their connections. In the text, each item is identified with the same letter as in fig. 4.1.

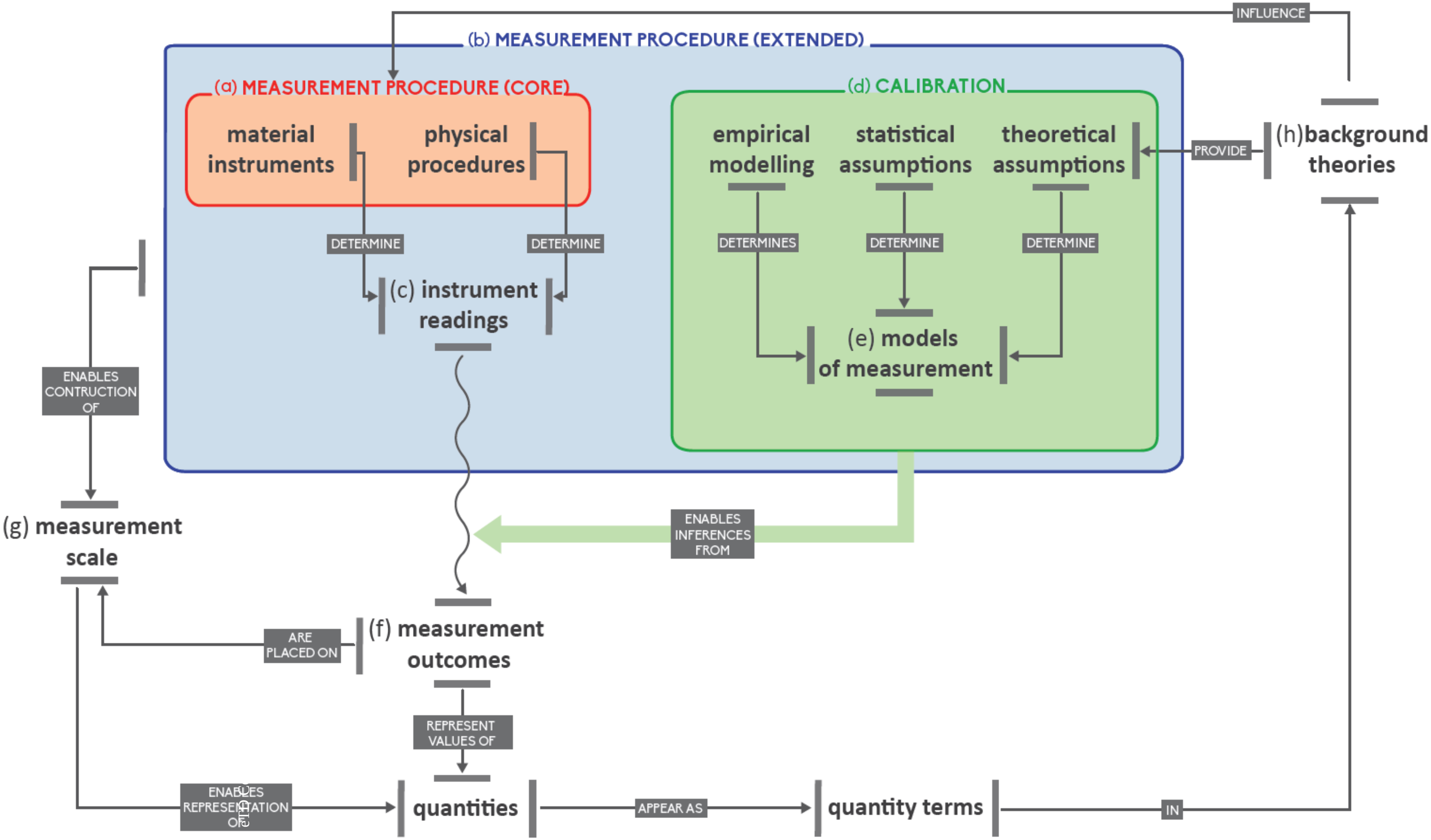


fig. 4.1: A representation of epistemic components typically involved in the process of measurement coordination

Measurement procedures are physical interactions between one or more epistemic subjects, a material apparatus, and some phenomenon occurring in an environment, as when we observe the mercury dilate in the column of a thermometer hanging on the wall. At the same time, the epistemic subjects purport to represent a certain relationship between quantities by means of the physical process taking place during the measurement interaction. In the case just mentioned, this happens when we read a measurement of temperature out of an indication of the length reached by the mercury in the thermometer column. In this sense, we can distinguish between two aspects of measurement. Measurement *senso strictu*, what I call 'measurement procedure core' (box (a) in the graph), can be viewed as the set of physical procedures and material instruments used for enacting a measurement procedure and, in some cases, (re)producing a phenomenon. What I labelled 'measurement procedure extended' (box (b) in the graph), is measurement in its broader sense, inclusive of its representational character and, therefore, of the host of inferential assumptions involved by its representational use. Central to this distinction is the difference between instrument readings and measurement outcomes.[101]

Instrument readings (item (c) in the graph) are the observations of the states of the material instrument measuring a phenomenon, once the physical process enacted during the measurement procedure has come to its end-state. For instance, when we place a mercury thermometer under our armpit, we must wait for a certain amount of time until the mercury in the thermometer column has expanded according to our body temperature. The end-state of the physical process, in this case, is when the mercury stops expanding, whereas the instrument reading is the length reached by the mercury column.

To understand how to construct and successfully perform a physical measurement procedure, the procedure itself is subject to calibration (box (d) in the graph). Calibration is the process through which models of the measurement procedure are constructed and tested, by modelling uncertainties and systematic errors of the procedure (or across procedures measuring the same quantity) under idealised statistical and theoretical assumptions (Boumans 2007, Frigerio et al. 2010, Mari 2003, Tal 2017c). The aim of calibration is to account for (ideally) all possible sources of measurement error given the best standards of precision available and, therefore, to improve the accuracy of a

[101] Cf. Tal (2017b), especially pp. 235-236, for a very clear exposition. As I mentioned above, the importance of this conceptual difference for the understanding of measurement coordination is stressed already by van Fraassen (2008). Cf. also Giordani & Mari (2012), and Tal (2013).

measurement procedure.[102] The outcome of the calibration process, that is, an explicit model of measurement (item (e) in the graph) or a less integrated set of assumptions, enables inferences from instrument readings to measurement outcomes, thus having a crucial role in the coordination of quantity terms with empirical content (Tal 2011, McClimans et al. 2017). However, it is rare that calibration fully precedes the performing of measurement, since the two are themselves subject to cycles of mutual refinement.

Given what I just mentioned, it follows that measurement outcomes are, in part, the product of a modelling process, i.e., calibration, which has as its object a certain measurement procedure. In other words, measurement outcomes (item (f) in the graph) are "the best predictors of the observed end-states of a measurement process relative to a particular theoretical and statistical model of that process" (Tal 2016a: 5). By modelling possible measurement errors and other confounding factors, and by recourse to statistical and theoretical assumptions, measurement outcomes are inferred from certain instrument readings. This means that some of the content of measurement outcomes is imposed by adjusting inconsistent observations in the light of idealised background assumptions. When measuring temperature with a mercury thermometer, an outcome of 37.5 °C is not simply the result of observing the instrument reading (i.e., the length of the mercury column). Inferring this outcome from the instrument reading presupposes an already constructed measurement scale (item (g) in the graph), the identification of a function relating the measured quantity and the quantity by means of which the former is represented, the modelling of the possible measurement errors, the reduction of the statistical relevance of confounding factors, etc.

By now, it should be clear that the process of achieving a form of measurement coordination, by which quantity terms acquire meaning, is influenced by several theoretical assumptions with different roles at different epistemic stages. Importantly, the theoretical background (item (h) in the graph) available to the epistemic subjects performing measurement and calibration often provides them with measurement laws that are crucial to anchor the calibration of a measurement procedure and, therefore, to enable inferences from instrument readings to measurement outcomes. Yet, theoretical background comes in also in the form of idealising statistical assumptions, which provide justification for the construction of measurement scales, that is, those mathematical structures which enable the

[102] As Tal (2017c: 34) points out, the term 'calibration' is often used with reference to "the empirical activity of detecting correlations among the indications of instruments, or between the indications of an instrument and a set of reference systems that are associated with fixed values. Values are then assigned to the indications of the instrument being calibrated so as to match previously known values, often along with a rule for extrapolating between (and beyond) those known values". This understanding mainly accounts for the practice of instrument making, which is only one aspect of calibration.

representation of measurement outcomes according to certain ordering, difference, and ratio relations.[103] However, the multiple roles of theoretical background should not lead to underemphasise the centrality of pragmatic considerations, empirical testing, and material instrumentation during calibration, especially in those epistemic contexts in which a sound theoretical understanding of the measurement interaction is lacking.

In the next section, I will apply the concepts just presented to my analysis of Ohm's experimental work on electric conductivity. These classifications will be a helpful guide to disentangle the measurement coordination between the quantity of exciting force and the measurement procedure deployed by Ohm. More specifically, I will show how analysing the details of the calibration process is necessary to pinpoint which assumptions were crucial to Ohm's specific contribution to this form of measurement coordination. This, in turn, will allow us to understand the role of measurement coordination with respect to the general coordination between Ohm's famous law and the empirical regularity that it expresses.

4.2 Measurement and coordination in Ohm's scientific inquiry

The long-standing achievement of Ohm's researches on electrical conductivity is the famous law named after him, which relates intensity of current (I) to voltage, or potential differential in modern terms (V), and resistance (R) according to the following equation:

$$I = V/R$$

Despite the widespread use of his law till nowadays, Ohm's own scientific endeavours have rarely been considered in detail by philosophers.[104] On the other hand, historians have shown a lot of interest in them, especially in the attempt to understand the reasons of the late reception of Ohm's work by the scientific community.[105] Yet, the fact that most historical analyses focusing on Ohm's investigations are all several decades old could be interpreted as an apparent lack of interest also from

103 Cf. Tal (2018) for a full exposition of the conventional aspects involved by the construction of measurement scales. It is particularly metric conventions, i.e., the conventions fixing the criteria for the equality of intervals of a quantity, that are crucial for determining the ordering of measurement outcomes and, thus, for achieving measurement coordination between the empirical results of measurement and their representation on a mathematical scale.

104 A notable exception is Heidelberger (1980).

105 Among the possible causes of Ohm's late reception, historians have identified the widespread impression that Ohm's law resulted from a purely theoretical deduction (Shedd & Hershey 1913), the hostility of the dominant Hegelian philosophy towards rigorous empirical researches (Winter 1944), Ohm's unorthodox use of the notion of tension (Schagrin 1963), and the highly mathematical character of his treatment, compared to the standards of contemporary German science (Caneva 1978).

contemporary historians in reconsidering the historical and epistemological import of Ohm's contribution.

One point on which most historians agree is that Ohm's work on conductivity and resistance introduced important elements of discontinuity in the scientific inquiry on the electric current, especially in the German context. Ohm has been viewed as a key figure anticipating a new wave of researchers that revolutionised the epistemic standards and values of German electrical science by their use of rigorous mathematical tools and their search for precise quantification (Caneva 1978). What many historical reconstructions emphasise as a central aspect of Ohm's work is his unorthodox use of some notions with respect to how electrical phenomena were conceived at his time (Archibald 1988, Atherton 1986, Gupta 1980, Schagrin 1963). More specifically, these authors focus on how Ohm's concept of *tension* (*Spannung*), characterised in purely geometrical terms in Ohm's (1827/1891) mathematical derivation of his law, was different from Ampère's, which was used to frame phenomena due to static electricity rather than due to current. According to these authors, Ohm's use of that concept for theorising required the abandonment of certain theoretical commitments and caused a shift in the inquiry on electrical phenomena.[106]

In this section, I reconstruct the experimental practice through which Ohm obtained his empirical results on conductivity and resistance. To do that, I will use a variety of historical sources, including Ohm's own original laboratory notes. In addition to that, I will use the categories from recent epistemology of measurement that I introduced in the previous section, as well as my own distinction between two kinds or scopes of coordination. My reconstruction aims at pinpointing Ohm's contribution to measurement coordination between the quantity of current intensity – 'exciting force', in his terminology – and the procedure to measure it. Achieving this coordination was crucial to the definition of what resistance was and it was important for Ohm to establish his famous law.

This case study provides support for two philosophical claims. The first claim is that the distinction between measurement coordination and general coordination is helpful to epistemological analyses, despite it being underemphasised by current literature. My analysis of the case study will show that the measurement coordination between exciting force and the procedure to measure it was not sufficient to achieve general coordination between Ohm's law and the relationship between current and resistance that it represents, since Ohm required an additional theoretical step for this general

[106] Cf. also Kuhn (1970: 469, footnote 14).

coordination to be justified. Therefore, it proves that a distinction between measurement coordination and general coordination can be crucial to epistemological analyses of specific scientific inquiries.

The second claim is that the details of a measurement procedure, particularly of the calibration process, can be crucial to understand measurement coordination and its relationship with general coordination. After analysing the details of Ohm's measurement procedures in two different sets of experiments, I will suggest that his trust in the instrument producing the current, i.e., the thermocouple, was an important source of justification for his measurement procedure and, therefore, it was an essential component of the measurement coordination required by Ohm to establish his law. On the one hand, this claim aims at counterbalancing the historical perspective according to which the main point to understand Ohm's establishment of his law is his conception of tension. What I suggest is that this view downplays the intricacies of measurement procedures and the role of material conditions in Ohm's advancements. Instead, I argue that we can understand his use of the term 'tension' only from within the context of new measuring techniques and their relationship with the rest of Ohm's conceptual apparatus. On the other hand, recognising the importance of the thermocouple for Ohm's measurement coordination provides an excellent starting point for my systematic discussion of the constitutive character of instruments and their materiality in section 4.3.

4.2.1 Background to Ohm's inquiry

4.2.1.1 Electrostatic concepts and the electric current by the early 1820s

By the time Ohm started his researches, theorising on static electricity had reached a relatively mature stage, while the inquiry on current phenomena – boosted by Volta's invention of the pile in 1800 and by Ørsted's discovery of electromagnetism in 1820 – was characterised by discussions on the appropriate concepts to be used for theoretical purposes. It is impossible to summarise in a few lines the intricate path towards the emergence of precise quantity concepts in electrostatics.[107] Yet, the influence of electrostatics on the understanding of current phenomena is relevant to the context in which Ohm operated. Therefore, I will briefly introduce the origin of some electrostatic concepts, as well as their relationship with how the inquiry on current phenomena was pursued before Ohm.

Towards the end of the 1780s Volta used the term 'tension' to represent the forces exerted by each point of an electrified body to free itself of its electricity and communicate it to other bodies.[108] Between 1796 and 1797 he made a terminological distinction between 'tension' and 'quantity of

[107] Cf. Heilbron (1979), especially chapter XIX, for an excellent summary.
[108] Cf. Heilbron (1979), especially p. 454, and references therein.

electricity' by suggesting that tension was influenced by the "mutual influence of the atmospheres" between different conductors, where by 'atmosphere' he referred to the power of different metals to 'push' the electric fluid (Volta 1918: 475).[109] Around the same time, quantities measuring the electrical action, that is, the strength of the force producing static phenomena, could be defined more precisely. Thanks to a powerful apparatus, based on his torsion studies, Coulomb managed to empirically establish the inverse law of electrostatics, by which the reciprocal attraction of negative and positive charges is in the inverse proportion to the square of their distances.[110] The term 'electrostatic force' identified the effects resulting from the attraction or repulsion of different opposite charges.

To explain the workings of his newly invented pile, Volta (1800) used the same distinction between 'tension' and 'quantity of electricity' that he developed during the 1770s and '80s. Whereas 'tension' identified a weak force responsible for static effects, 'quantity of electricity' referred to the amount of fluid – set in motion by the contact of different metals – that was producing an electromotive force causing the current to flow. Therefore, this distinction was crucial to the differentiation between the new current phenomena and the known static ones.

Yet, the influence of electrostatic concepts in framing the understanding of current phenomena was a lasting one (Brown 1969). The electrostatic theory of the pile was popular until 1820 when, after Ørsted's discovery of electromagnetism, Ampère introduced some terminological innovation to the study of electric current.[111] With 'electromotive action', Ampère (1820) referred to the phenomenon taking place inside the pile, which produced an 'electromotive force'. This force caused the current to flow within a closed circuit, or it generated a tension if the terminals of the circuit were open.[112] Furthermore, in 1823 Ampère established a force law stating that the force of attraction or repulsion between two charged wires is proportional to their lengths and to the intensity of their charge. This regularity subsequently proved to be central to the understanding of electrical phenomena in terms of their magnetic effects.

109 Cf. Pancaldi, (1990): 128-133.

110 Gillmor (1972) provides an excellent account of Coulomb's scientific work. See below, section 4.2.2 for more on Coulomb's torsion studies and their relevance for the study of electricity. Cf. also Heilbron (1979), pp. 458-473 for earlier attempts at quantifying electrical action and antecedents to Coulomb's law.

111 Cf. Steinle (2016, especially ch. 1-4) for the best account to date of Ørsted's path to the discovery of electromagnetism and Ampère's conceptual innovations.

112 Cf. also Caneva (1980). Brown (1969, pp. 84-85) emphasises that Ampère explicitly refers to Volta's distinction between tension and current effects.

4.2.1.2 Measuring conductivity and resistance before Ohm

Experiments on conductivity and resistance had been performed well before the invention of the pile. In 1753, Beccaria observed that an increase in the cross-section of a tube filled with water was correlated to a more powerful shock caused by a discharge passing through the tube. Already by 1775 Cavendish had measured through direct sensation the differential in electric charge between glass tubes of different widths filled with a salt solution.[113] His method to determine a difference of discharge consisted in placing his own body in the circuit with the test-tube, and then to judge by the sensation experienced in comparison with a standard test-tube.

Many advances before Ohm's researches brought more precision and reliability to the measuring techniques and introduced some conceptual refinements. Priestley used a more reliable measurement procedure than Cavendish to measure discharge. He measured the length of the maximum air gap across which he could just get a spark to jump when placing different resistors in circuit: the larger the gap, the better the conductor. The impossibility of experimenting with a constant current, which became available only after the invention of the pile, was a substantial limitation to empirical researches on the relationship between current and resistance, since it was impossible to test for non-instantaneous effects of resistance. Davy was plausibly the first experimenter to use a steady current and found that wires having the same ratio of length to cross-section had the same resistance. Becquerel established, also by experimenting with a constant current, that the conducting power varied inversely as the length of the conductor, and he determined experimentally that the total current is the same in every part of a series circuit.[114]

4.2.2 Ohm's experimental work (I)

When he started investigating galvanic phenomena, Ohm's aim was that of discovering the law by which metals conduct electricity.[115] In a first set of experiments, which he reports in his (1825a/1892), Ohm attempted to establish how the conductivity of an object is affected by its length, while in a second set, which provided part of the material for his (1825b/1892) and (1826/1892), he attempted to establish the relative conductivity of different metals. His apparatus for these two sets of experiments consisted of a wet cell – a voltaic battery – as the source of electricity, to which different

[113] Cf. Heilbron (1979, pp. 477-89) for a well-rounded exposition of Cavendish's quantitative concepts and measurement techniques.
[114] This fact had a fundamental importance in Ohm's theoretical treatment. Cf. below, section 4.2.4.
[115] My reconstruction of Ohm's experimental work draws from Appleyard (1930), Gupta (1980), McKnight (1967), Schagrin (1963), Shedd & Hershey (1913), and my own research in the Ohm Collection at the Deutsches Museum Library Archives in Munich.

conductors could be connected as part of an external circuit, and a torsion balance measuring the impact of the length of a certain conductor in the electric circuit.

The balance, built by Ohm himself, consisted of a magnetic needle suspended over the conductor from a support. When deflected, the needle was brought back to the '0' position by turning of the support, and torsion of the suspension was, therefore, taken as the measure of magnetic force. For this measuring technique Ohm was relying on the foundational work on the theory of torsion by Coulomb. From the late 1770s, Coulomb started developing a theory of torsion in thin threads, and he showed how the torsion suspension could provide a method to measure small forces with great accuracy (Gillmor 1972, ch. 5). By means of careful experimentation, in the 1780s Coulomb identified a law of torsion, relating the torque (the momentum of the torsion force), the angle of torsion, the diameter and length of the thread or wire, a constant rigidity coefficient. Knowledge of torsion laws was, therefore, crucial to build precise torsion balances that could be used for accurate measurements of magnetic and electric phenomena.[116]

In the first set of experiments, Ohm placed wires of different lengths and same cross-section between the terminals of the circuit, to test for the effects of differential length against a very short 'standard' wire. While length was his independent variable, the dependent variable under test was the loss of force (*Kraftverlust*) of the circuit. The quantity term 'loss of force' was operationalised as the reduction in torsion of the magnetic needle while a test wire was in the circuit, compared to that effected when the standard wire was placed in the circuit. Ohm's choice of loss of force as the variable to test is not surprising, since by then it was known that a conductor diminished the magnetic effect of an electric charge inversely to its conductivity.[117] The major material difficulty in this procedure was the polarisation of the wet cell, which caused sudden surges of electricity when the circuit was opened or closed to replace a test wire. Ohm coped with this problem by waiting to observe the instrument reading until the magnetic needle stopped moving, rather than immediately. Another issue was the progressive decrease in the intensity of the electric stream produced by the wet cell. This and other complications provide interesting elements to assess the methods used by Ohm to account for measurement error and, thus, to infer certain measurement outcomes from his instrument readings.

[116] See Gillmor (1972, ch. 6) for more on Coulomb's application of his torsion studies and of his balance to his work on magnetism and electrostatics.

[117] The fact that the poorer the conductor the less the magnetic effect of the electric stream was known at Ohm's time, but none of the many explanations suggested for it had reached a consensus. Schagrin (1963: 541, footnote 21), for instance, refers to J. B. Biot's *Précis Éleméntaire de Physique* (Deterville, Paris 1821), 2nd ed. as one of Ohm's possible sources. This phenomenon is in accordance also with Ampère's force law.

As it can be seen in Ohm's own laboratory notes, before experimenting on the correlation between loss of force and length of the conductor, he made experiments to study the rate of decay of the wet cell.[118] These observations were apparently repeated more times, even after the first sets of experiments to relate length of the conductor and loss of force, as reported in the notes taken on the following days.[119] By establishing the rate of decay, he could model the measurement error due to the decreasing intensity of the electricity produced by the wet cell and factor it in to infer measurement outcomes of loss of force from the readings of the torsion balance. Unfortunately, from Ohm's laboratory notes there is no indication of what instrumental apparatus he used to measure the rate of decay. Nonetheless, it is quite plausible that he used the very same torsion balance by means of which he collected the measurements during the experiments on the effects of length on measurements of loss of force. In addition to the rate of decay, he modelled the internal resistance of the apparatus, not to be confused with the resistance due to the length of any specific conductor that he was testing, which he called 'reduced length'. He accounted for this source of error by repeating the experiment several times changing the composition and geometry of the setting, particularly the external circuit. Finally, he averaged the readings of the torsion of the needle when the standard conductor was in place in order to statistically reduce the measurement error resulting from the comparison between measurements of test wires and those of the standard conductor.

To sum up, the elements involved in Ohm's measurement practice in this first set of experiments are the following:

- *Material instruments*: torsion balance, wet cell, test wires.
- *Physical procedures*: placing different conductors in the open circuit, reading the indications of torsion balance.
- *Instrument readings*: different gradients of deflection of the magnetic needle in the circle of the torsion balance.
- *Calibration* (with the type of modelling involved indicated in bracket):
    1. Waiting to read the indication of the balance to account for the polarisation of the wet cell (methodological/pragmatic);
    2. Testing of the rate of decay of the wet cell (empirical);
    3. Repetition of the experiment with different geometrical compositions of the setup to model the internal resistance of the apparatus (empirical);

[118] Cf. Item 904 – NL267/017 of the Ohm Collection, 7th page.
[119] Cf. Item 904 – NL267/017 of the Ohm Collection, 8th and 9th page.

4. Averaging of the balance readings when the standard conductor was in place (statistical).

The results of this modelling process were factored in to infer measurement outcomes from instrument indications.

- *Measurement outcomes*: values of loss of force represented as the reduction in torsion of the magnetic needle when a test wire was in the circuit, compared to position of the needle when the standard conductor was in place. '0' value provided by the average of measurements with the standard conductor in place.
- *Quantity concept*: loss of force (*Kraftverlust*).
- *Resulting theoretical model*: logarithmic formula $v=m\log(1+x/a)$, where v is the loss of force, x is the length of the conductor, a and m empirical constants, the former depending on the internal resistance of the circuit, the latter depending on many factors, including a.

As this list shows, Ohm's modelling of the measurement error was mostly a matter of careful empirical testing and pragmatic considerations, and a substantial part of it aimed at accounting for the instability of the phenomenon on which he wanted to experiment, i.e., what we would call an electric current. On the contrary, with respect to his measurement instrument, i.e., the torsion balance, calibration was not necessary. As I mentioned above, the workings of the balance were underpinned by Coulomb's torsion law and, by the time Ohm started his inquiry, Coulomb's own experimental results provided sufficient justification to trust the precision and reliability of this instrument for the measurement of magnetic and electrostatic phenomena. Therefore, trust in Coulomb's torsion law was sufficient to rely on the capacity of the balance, which embodied that law, to provide sufficiently accurate measurements. The use of the balance to measure both the effects of placing different conductors in the circuit and to model the rate of decay of the wet cell would not imply a circularity of the kind described in section 4.1.2, since the torsion law provided independent justification for the use of the balance. At the same time, Ohm's trust in Ampère's force law, stating the proportionality between intensity of charge and magnetic effects, justified his inference from instrument readings of the balance, that measured magnetic effects, to measurement outcomes of loss of force.

Since Ohm's measurement procedure relied on the sound justification provided by Coulomb's torsion law and by Ampère's force law, one might wonder what prevented him from measuring the force of the circuit directly, rather than measuring its loss of force due to each test conductor compared to that due to the standard wire. In other words, we might ask what led him to consider the variable of loss of force as a safer choice, apart from his knowledge of the fact that a conductor diminished the magnetic effect of an electric charge inversely to its conductivity.[120] The analysis of the physical procedure and of the assumptions underlying it suggests that the strongest limitations to the reliability of Ohm's measurement procedure came from his productive instrument, i.e., the instrument responsible for producing the phenomenon on which to experiment (cf. Heidelberger 2003). Although Ohm carefully attempted to account for the instability of the electricity produced by the wet cell, it seems plausible that his lack of confidence in the stability of the phenomenon influenced his choice of the experimental variable. Measuring the loss of force of the circuit when each test conductor was in place in comparison to the circuit with the standard conductor may have also seemed a better choice in the light of the necessity to alternate the testing for loss of force with tests for the rate of decay of the wet cell, as it appears from his laboratory notes.

With respect to the empirical results, historians disagree on whether considering this first set of experiments as unsuccessful or taking the resulting logarithmic formula as a good approximation to Ohm's law. In addition to the issues highlighted above, his formula was not reliable beyond a certain length of the conductor, since Ohm himself (1825c/1892: 12) recognised that his approximations failed for very long conductors, that is, when the external resistance approaches or exceeds the internal resistance. This demonstrates also the limitations of the measurement scale that he developed to place the values of loss of force in relationship with one another, so as to account for the degree of difference between the measurement outcomes of loss of force. These limitations exhibit a typical issue within the process of measurement coordination, i.e., that after a certain point of a constructed measurement scale its structure is no longer reliable, since it does not support inferences from instrument readings to measurement outcomes, because the measuring device has not been – since often cannot be – robustly tested.[121]

4.2.3 Ohm's experimental work (II)

In January 1826, Ohm carried out another set of experiments, under Poggendorff's suggestion to repeat his first set by using a copper-bismuth thermocouple instead of a wet cell as a source of electricity (fig. 4.2). The junction points of this thermoelectric source were immersed in boiling water and melting ice, which were supposed to set a constant temperature differential across the thermocouple and, thus, produce a constant electric stream. The torsion balance was only slightly adapted to fit to the new setting.

[120] Cf. above, footnote 117.
[121] Cf. Chang (2004: Ch. 3) on similar issues with thermometry.

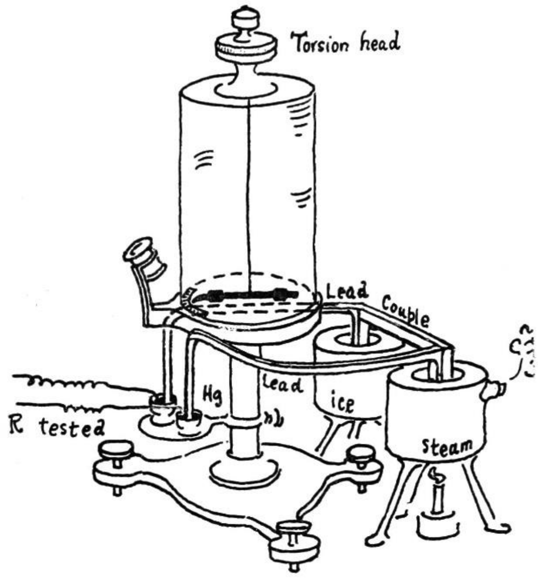


fig. 4.2: Ohm's torsion balance and thermocouple. Reproduced from Shedd & Hershey (1913: 607)

Ohm's laboratory notes present a seemingly smooth process of data gathering.[122] Three main points stand out in these notes:

1. The electric surges observed in the first set of experiments are now totally absent;
2. The deflection of the needle, measured by the turning of the divided circle, is taken as the measure of the intensity of the electric stream;
3. The standard conductor, crucial in the first set of experiments, is now absent.

In a note towards the bottom of the page, Ohm writes that the thermocouple eliminated the violent electric surges when the circuit was opened or closed, due to the high polarization of the wet cell. Despite the low internal resistance of the new apparatus, calibration as empirical modelling could not be entirely dispensed with. In the experimental trials conducted in the following days, Ohm seemingly attempted to minimise other sources of error and influences of independent variables, such as junction temperature deviations and contact resistance.[123]

Ohm's dependent variable, this time, was not the loss of force of the circuit. Instead, he reported the observed force of the electric stream for each test wire, what he named 'exciting force' (*erregenden Kraft*).[124] This change of variable is related to the absence of the standard wire, on the basis of which

[122] Cf. Item 904 – NL267/017 of the Ohm Collection, 25th page.
[123] Cf. Item 904 – NL267/017 of the Ohm Collection, pages 26th to 39th.
[124] This variable assumed the name of 'electromotive force' in Ohm's 1827 treatise, in conformity with the nomenclature introduced by Ampère.

he compared and inferred the outcomes of loss of force in his first set of experiments. In other words, instead of measuring the loss of force of the electric stream depending on the length of the different conductors compared to a standard conductor, Ohm measured directly the force of the electric stream for each conductor. What prompted this change of the measured quantity? First I must describe the outcome of Ohm's experiments. The final formula representing his results was:

$$X = a/(b+x),$$

where X is the strength of magnetic action (representing current strength), x the length of the wire, a the exciting force, and b the resistance of other parts of the circuit.

To understand the change of variable from loss of force to exciting force, it is crucial to consider the details of Ohm's new experimental setup. He knew the workings of the measuring instrument in the apparatus, the torsion balance, from the previous sets of experiments. In both settings, he used the balance to measure the impact of each conductor on the electric stream, where the intensity of the latter was measured by observing its magnetic effects, in accordance with Ampère's force law. The major change in the experimental setup was the thermocouple. The thermocouple had a productive role, since its purpose was that of producing a constant electric stream, with which Ohm could test the effects of placing wires of different lengths in the circuit. To do this, Ohm had to rely on the stability of the electric stream produced by the thermocouple, something which was lacking in his previous experimental setting. Ohm took the stability of the temperature at the endpoints of the thermocouple as a warrant for the intensity of the electric stream not decreasing across time. In the basic thermocouple setup, which I described in the beginning of this section, one endpoint was immersed in melting ice and the other in boiling water, thus setting a fixed temperature differential of 100 C°. The principle underlying the workings of the copper-bismuth thermocouple dated back to 1822, when Seebeck published the discovery of the so-called 'Seebeck effect', that is, the production of a thermomagnetic effect caused by a temperature difference between the ends of two wires of different metals joined at both ends.

Here is the list of elements involved in Ohm's measurement practice in this set of experiments:

- *Material instruments*: torsion balance, copper-bismuth thermocouple, test wires.
- *Physical procedures*: placing different conductors in the open circuit, reading the indications of torsion balance.

- *Instrument readings*: different gradients of deflection of the magnetic needle in the circle of the torsion balance.
- *Calibration* (with the type of modelling involved indicated in bracket):
    1. Testing for the junction temperature deviations of the thermocouple (empirical);
    2. Testing for contact resistance of the circuit (empirical).
- *Measurement outcomes*: "absolute" values of exciting force, with no need for a standard conductor to fix a '0' value; the values of exciting force can be placed on an interval scale, the metric convention being dependent on calibration assumptions and background theoretical commitments.
- *Quantity concept*: 'exciting force'.
- *Resulting theoretical model*: mathematical formula expressing the inverse proportionality between intensity of electricity and resistance, and direct proportionality between the former and exciting force ($X = a/(b+x)$ in Ohm's experimental paper, I =V/R in modern notation).

Within this experimental setting, fixing a stable temperature difference seemed to reliably ensure the constancy of the electric stream across time, which was a major problem in the first set of experiments. In 1823, Ørsted put forward an interpretation of the thermomagnetic effects caused by temperature difference in terms of thermoelectric effects.[125] It is on the basis of this recent and not uncontroversial interpretation that Ohm could consider the thermomagnetic effects produced by the thermocouple – and measured by his reliable torsion balance – as produced by a constant electric stream. In other words, the stability of the magnetic effects produced by the thermocouple at a set temperature differential ensured Ohm that the electric stream on which he was experimenting was a stable one.

However, Ohm could not obtain his law only by experimenting on an electric stream at a single fixed intensity, i.e., the one set by the 100 C° temperature differential. To test the robustness of his regularity – i.e., in modern terms, to test the invariance of the resistance coefficient independently of current strength – he had to produce electric streams of different intensities. Recall the final formulation of the regularity that he established: $X = a/(b+x)$. By repeating the experiment with different initial temperatures for the thermocouple, Ohm believed that he could determine that the value of b remained the same while that of a changed, $(b+x)$ being the total resistance of the circuit. To test for the effects of length of the conductors on the exciting force of the circuit, he had to make sure that variations in temperature difference without alterations of the resistance parameters, i.e., by keeping the circuit the same, should result in proportional variations in the magnetic effects measured by the

[125] Cf. Ørsted & Fourier (1823).

torsion balance. In other words, Ohm had to make sure that magnetic effects, taken as a measure of the intensity of the electric stream, were directly proportional to temperature difference, so that he could properly exclude that measurement outcomes of exciting force of the circuit were influenced by variables other than the length of the conductors placed in the circuit.

In his paper, Ohm (1826: 153) reports having carried out his experiment with *one* temperature differential that was not the same as the one in the basic setting.[126] According to Schagrin (1963), who offers the most detailed account to date of Ohm's experiments, Ohm also believed – in addition to Ørsted's thermoelectric interpretation – that Ampère and Becquerel showed the direct proportionality of tension and temperature differential between the contact points of an open thermoelectric source.[127] Yet, Ohm was experimenting on a *closed* circuit, and tension, as conceived by Ampère and Volta, referred to effects due to static electricity in an open circuit. This point leads Schagrin to argue that the use of the thermocouple was not a necessary factor in Ohm's decision to change the variable to measure from loss of force to exciting force, since he could have kept measuring loss of force even with the thermocouple setting. On the other hand, in his view, Ohm's use of 'tension' in a discontinuous way with respect to Ampère's was the crucial factor that enabled Ohm to achieve his law.

I propose a different interpretation than the one offered by Schagrin. The evidence from the discussion above seems to suggest that Ohm spent limited time in testing experimentally the proportionality between temperature differential and force of the thermoelectric circuit, and that he was quite ready to rely on others' accounts of it, despite the novelty of the phenomenon, the early stage of its theoretical understanding, and the incoherence between his conceptual apparatus and those of others. This, in turn, seems to suggest that Ohm's trust in the reliability of the thermocouple did not require a full theoretical, before than experimental, grasp. Certainly, contrary to the torsion balance, the theoretical understanding of the workings of the thermocouple was still limited. Indeed, he *could* have kept measuring loss of force even with the thermocouple setting. However, Ohm deemed the thermocouple much more reliable than the wet cell for its capacity of producing a stable electrical stream (Heidelberger 2003). Therefore, Ohm conducted the empirical calibration of the thermocouple, as part of modelling of his measurement procedure only to a limited extent. His trust in the *material* reliability of his productive instrument in (re)producing a phenomenon was *sufficiently*

[126] Unfortunately, the difficulty in interpreting his laboratory notes did not allow me to check whether he reported experiments made with further temperature differentials.
[127] Cf. Schagrin (1963: 545, especially footnote 27) and references therein.

coherent with available knowledge of thermo-magnetism, with his trust in the accuracy of the available thermometric standards – by which he measured the temperature differences while experimenting,[128] – and with the measurement laws involved by his measurement procedure (Ampère's force law and Coulomb's torsion law), so that he felt justified to measure exciting force instead of loss of force and, thus, rely on a different form of measurement coordination.[129]

4.2.4 From measurement coordination to general coordination: 'tension' in *Die Galvanische Kette*

When introducing the case of Ohm's inquiry, I emphasised that many historical accounts consider Ohm's use of the term 'tension' (*Spannung*), characterised in geometrical terms in his mathematical treatise, as discontinuous with respect to Ampère's. As I explained in section 4.2.1, Ampère distinguished between effects depending on tension and effects depending on current. In Ampère's view, tension disappeared at the closing of the circuit because no electrostatic action could be detected. Therefore, the notion of tension only made sense if related to an open circuit (e.g., the endpoints of an open pile), whereas 'current' referred to phenomena taking place in a closed system. At the end of section 4.2.3, I discussed Schagrin's view that Ohm's unorthodox use of 'tension' with reference to his closed thermoelectric circuit was the crucial factor that enabled Ohm to achieve his law, and I suggested a different interpretation. In this section, I want to put forward an epistemological interpretation of one of Ohm's assumptions concerning his notion of tension. Ohm's assumption that there is tension between every two points of an open or closed circuit, although it had an empirical basis, acquired the status of a deductive axiom in Ohm's mathematical treatment. What I suggest is that, even if Ohm made this assumption already during the experimental phase, it is best understood as providing justification for the general coordination between Ohm's law and the empirical regularity that it describes, rather than for the measurement coordination between exciting force and the procedure to measure it.

Ohm conceived of the electric flow as of a modification in the spatial distribution of exciting force in a conductor, modification produced by the attempt of the system of conductors to reach an equilibrium state. In his view, the difference in exciting force between every two parts of a conductor, what Ohm referred to as 'tension', generated the flow of electricity. Ohm seemingly characterised tension as equally distributed across the electric circuit and not vanishing once the circuit was opened,

128 At the time of Ohm's experiments, controversies on the 'correct' thermometric standard had not yet come to an end (cf. Chang 2004: ch. 2 and 4).
129 Interestingly, Fechner went back to using a wet cell as a productive instrument to test experimentally Ohm's law (cf. Caneva 1978: 110-111).

so that his quantity term 'exciting force' could keep referring to the electricity flowing in the thermoelectric circuit independently of the presence of a conductor between the endpoints of the circuit.[130] In this sense, the assumption that tension was equally distributed in the circuit was already presupposed, or embedded, in his notion of exciting force. This assumption helped Ohm making sense of his measurement practice, which was based on the continuous opening and closing of the thermoelectric circuit to replace the conducting wires, in order to test for the effects of their length on exciting force.

In *Die Galvanische Kette matematisch bearbeitet* (1827/1891) the notion of tension became even more central to Ohm's mathematical derivation of the law. Ohm developed his mathematical treatment by imagining a homogenous and uniformly thick ring, in which any two points are characterised by one and the same tension.[131] Ohm characterises tension as the difference in exciting force (which he now calls 'electromotive force') between every two parts of a conductor. As I mentioned above, this difference causes the electricity to flow, since the circuit attempts to reach a state of equilibrium. In other words, an unaltered state of tension between every two parts of the conductor causes the constant transmission of electricity. This statement seems justified by the fact that "each particle of the conducting medium situated in the circuit of action receives each moment just the same amount of the transmitted electricity from the one side as it gives off to the other, and therefore constantly retains an unchanged quantity" (Ohm 1827/1891: 22). That the total electricity is the same in every part of a circuit had been established experimentally by Becquerel in 1825, and it was a fact well known by Ohm, who acknowledges Becquerel's empirical discovery (Ohm 1827/1891: 67). Even though this was an empirical fact, Ohm constructed an axiom out of it, he 'elevated' it to a fundamental principle (Archibald 1988: 145). According to this axiom, tension is between all points of an open or closed circuit, thus characterising it, by definition, as constantly distributed along the circuit. If the electricity flowing through the circuit was the same everywhere (given Becquerel's discovery), and it was transmitted only between adjacent particles, as postulated by Ohm, tension had to be constantly distributed along the length of the conductor, and Ohm represented it as a straight line in the geometrical representation of his law. In addition, this was independent, in principle, from the fact that the circuit may be open or closed (Ohm 1827/1891: 65-66).

130 In some historical accounts (e.g., Archibald 1988) the two terms 'tension' and 'exciting force' are sometimes identified. The translation of Ohm's 1827 treatise itself translates *Spannung* as 'electroscopic force' (p. 14, footnote 1). However, this choice obscures the distinct roles of these terms in Ohm's conceptual apparatus.

131 Further idealising assumptions to model the abstract galvanic circuit were the mono-dimensionality of the propagation of electricity and the assumption that galvanic phenomena do not vary with time, given the constancy in time of the sources of current.

In sum, Ohm's assumption of tension as equally distributed between every two points of a circuit was based on Becquerel's empirical discovery that the total electricity is the same in every part of the circuit. Ohm made this assumption already during the experimental phase to make sense of his measurement outcomes of exciting force. Yet, the assumption of equal distribution of tension acquired the status of a deductive axiom in his mathematical treatise, in virtue of Ohm's definition of tension as the difference of exciting (or electromotive) force between every two points of a conductor. This was a distinctively theoretical assumption that provided part of the justification for representing the empirical relationship between electromotive force and resistance in terms of his law. Therefore, it is best understood as an assumption involved in the general coordination between Ohm's law and the regularity it describes, rather than of the measurement coordination between exciting (or electromotive) force and the procedure to measure it.

4.2.5 What is constitutive in Ohm's scientific inquiry?

Now that I have disentangled the two related kinds – or, better, scopes – of coordination in Ohm's scientific inquiry on current and resistance, I can turn to identify what epistemic components can be viewed as constitutive, and to what degree, in these forms of coordination. Let us start from the end. As I have emphasised in the last section, Ohm's theoretical assumption of equal distribution of tension was based on Becquerel's empirical discovery that the total current is the same in every part of the circuit. Yet, Ohm turned this empirical fact into an axiom and defined 'tension' as being constantly distributed between all points of an open or closed circuit. In other words, this empirical fact was transformed into a postulate that acquired a new epistemic status and became a foundational building block for the formulation of Ohm's law. From the epistemological point of view, this dynamic resonates Poincaré's (1902) discussion of how certain principles of mechanics are definitions 'in disguise', since they have their origins in experimental regularities, but are elevated to the state of fundamental axioms with an absolute value. Friedman (2009) reinterprets Poincaré to suggest that, in those cases, the new epistemic status of these principles makes them constitutive of empirical regularities, as in the case – he suggests – of the principle of relativity and the light principle in Einstein's special theory of relativity.[132] In my view, this interpretation is helpful to understand what happened with Ohm's notion of tension, when Ohm transformed Becquerel's inductive generalisation in a deductive axiom to provide a definition of 'tension' and, therefore, justify the coordination between Ohm's formulation of his law and the empirical relationship between current and resistance. On the lines of the cases I discussed in chapters 2 and 3, this switch in the epistemic function of a

[132] Cf. above, chapter 2, section 2.2.2 for more details on Friedman's example of the light principle.

theoretical principle from inductive generalisation to deductive axiom is expressed by an increase of all three feature of constitutivity, quasi-axiomaticity, generative potential, and empirical shielding, of the principle that current is constantly distributed among all points of a circuit, independently of the latter being open or closed.

However, theory does not take up a justificatory role only so late in the Ohm's inquiry. As we saw, the modelling of the measurement process, i.e., calibration, also played a crucial epistemic role in the process of measurement coordination. In both sets of experiments, measurement outcomes were not merely the end-products of a purely physical interaction between measuring instrument and phenomena. Coulomb's torsion law underlay the working of Ohm's measurement instrument, that is, the torsion balance, whereas Ampère's force law enabled him to read measurement outcomes of loss of force in the first set, and of exciting force in the second set, out of instrument readings of magnetic strength. Therefore, these independent empirical regularities, mathematically expressed, played an important justificatory role already in measurement coordination. I will come back again to the specific contribution of these epistemic components in the next section.

After discussing the role of theoretical principles in general coordination and in measurement coordination, one question still stands out. Is there a specific epistemic contribution of the *material components* of the measurement procedure to the constitution of quantities in measurement coordination? On the one hand, the core of the measurement procedure appears to frame a material space of possibilities. In other words, instrument readings seem to embody certain constraints coming from the material structure of the measurement interaction, that is, from the relationship between the measuring instrument, the epistemic subjects, and the environment. On the other, the material apparatus often embodies measurement laws that seems to justify theoretically the measurement interaction. Therefore, *prima facie*, it might seem that the theoretical background, in the form of measurement laws, carries most of the justificatory burden in measurement coordination. However, in my view, this is not the whole story. In some circumstances, which I discuss in the next section, scientific instruments performing specific epistemic roles – as in the case of Ohm's thermocouple – can have, to some degree, a constitutive character in measurement coordination.

4.3 The constitutive character of scientific instruments and their materiality

4.3.1 The epistemic roles of instruments and their materiality in scientific inquiry

In this final section, I will focus on the epistemic role of scientific instruments with respect to their involvement in coordination. Scientific instruments (together with physical models and laboratory settings) are so obvious that they are often overlooked by philosophical analyses of science. However, not only do they comprise the material interface between the epistemic subjects and the world from which they gather empirical input, but they also embody certain limits of our scientific inquiries that are different from those coming from our theoretical representational resources. Yet, the distinctive role of the *materiality* of scientific instruments in experimentation has proved hard to pinpoint. Morgan (2003) develops a continuum where the extent to which materiality (especially, but not only, of scientific instruments) is crucial to experimentation depends on the type of experiment – virtual or physical – to be performed. Here, I develop Morgan's approach to show that the relevance of materiality comes in degrees even in the apparently more straightforward case of real laboratory experiments. What I want to suggest is that the relevance of materiality depends, in part, on the experimental role of instruments. On the one hand, the material character of instruments constrains the experimental manipulation and, therefore, it impacts the epistemic outcome of experimentation in a different way from theoretical constraints. On the other, the breadth of this impact does not depend only on the role that the apparatus plays within the experimental context, but also on the extent to which the representational character of the experimental setting is justified by theoretical background.[133] In what follows, I elaborate my view of materiality and, in the beginning of the next section, I bring back the discussion concerning the relationship between theory and experimentation that I developed in chapter 1. Both are required to understand the constitutive character of scientific instruments in coordination.

Materiality is an aspect quite underestimated by contemporary philosophy of science despite its pervasiveness in scientific inquiry. The literature agrees on acknowledging a wealth of epistemological roles played by the material components of scientific experimentation, and most of the effort has been channelled into attempts at classifying scientific instruments accordingly. To this purpose, several taxonomies have been suggested as a starting point for analysing the scientific and philosophical importance of instrumentation (cf. Baird 2003, 2004; Boon 2004; Harré 2003; Heidelberger 2003). In what follows, I will freely draw on categorisations put forward by these authors to construct a novel one. The latter will be based on the principle that, depending on the epistemic role of instruments, their materiality constrains the space of possibilities of experimental manipulation in a certain way, thus determining certain features of the experimental outcome.

[133] See above my discussion of theory-dependence and the relationship between theory and experimentation in Chapter 1, section 1.3.3.

I adopt three main groupings of instruments with respect to their epistemic function: productive, modelling, and quantifying instruments.[134] These groupings sit on a continuum and are ordered from left to right depending on the extent to which material constraints from instrumentation determine features of the experimental manipulation and, consequently, features of the result (e.g., a phenomenon reproduced in the laboratory, a measurement outcome, etc.) of the experimental interaction (fig. 4.3). However, I also distinguish between subclasses of instruments, depending on certain structural features. Some of the structural groupings fall clearly within one of the functional groupings, some others span over more functional groupings, and the subclasses I introduce are not an exhaustive list. In addition, instruments with different functions can be combined in a single apparatus (as in the case of Ohm's experiments, which I will go back to in the next section) and it is possible that the groupings are not mutually exclusive.

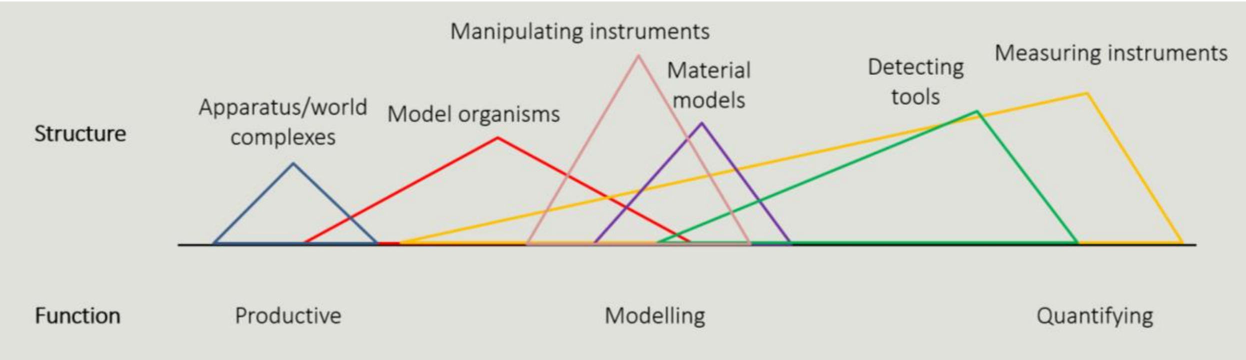


fig. 4.3: Continuum of scientific instruments in real experiments, classified according to their epistemic function (bottom) and common structural features (top). The groupings are ordered according to the extent to which the material features of experimentation determine and justify features of the outcome of experimentation (left = maximal extent; right = minimal extent)

Functional groupings

i. *Productive instruments*: instruments belonging to this grouping contribute to the *production* of phenomena that are not present as such in nature. In this case, the material features of the instrument are maximally relevant to the representational features of the experimental outcome, because the phenomenon under investigation is *ontologically* inseparable from the instruments.[135] In fact, experimentation that deploys instruments with a productive function

[134] My functional groupings reflect, by and large, those suggested by Boon (2004, 2015).

[135] The productive character of some kinds of instruments should not be confused with the general 'constructive' character of instrumentation, understood as the influence of instrumentation in determining features of the epistemic outcome of scientific inquiry. In my view *all* instruments are to be considered, to various degrees, as constructive because of their function of sensory and cognitive narrowing (cf. Gooding 2003). What makes them constitutive, independently from theory, is the extent to which the narrowing is *justified* by material features of the apparatus rather than by theory.

is often characterised by a blurring of the sharp distinction between object or phenomenon under inquiry and apparatus (Boon 2004). Consequently, material features of these instruments are inextricably bound to the way the phenomena are generated, reproduced, and manipulated, thus determining most features of the outcome of the experimental interaction.

ii. *Modelling instruments:* instruments belonging to this grouping interact with the phenomenon under inquiry by constraining it, modifying it, or reproducing it within an artificial experimental set-up in order to simplify it, usually for the purpose of experimental manipulation. In this case, the material features of instrumentation do not contribute to the *production* of the phenomenon of interest, even if they interact with the phenomenon on an ontological (causal) level, and the apparatus can usually be clearly distinguished from the phenomenon. Still, the apparatus has an impact on the character of the experimental manipulation, because the possibilities of manipulating the phenomenon of interest are partly constrained by the specific material features of the instrumentation.

iii. *Quantifying instruments*: instruments belonging to this grouping aim at representing one or more properties of the phenomenon under investigation in quantitative terms. Quantifying instruments certainly interact causally with the object of inquiry, in order to represent one or more of its features. Yet, as in the case of modelling instruments, they are not supposed to contribute to the production of the phenomenon to be quantified, which is usually identified *qua* phenomenon prior to the attempt at quantifying any of its properties. Still, their material features have an impact in determining the outcome of the experimental interaction, because the material form of the instrument embodies a field of possibilities that guides and constrains the possible transformations that eventually represent the outcomes of measurement (Baird 2003, Mets 2019). However, the material features of instrumentation, in this case, constrain actual experimental manipulation to a lesser extent than modelling instruments, since these constraints impact the possibilities of representing certain attributes, rather than the intervention on them. Therefore, this grouping is located on the right end of the continuum.

<u>Structural subclasses</u>

a. *Apparatus/world complexes*: Harré (2003) uses this notion with reference to those instruments that bring about phenomena that do not exist as such in nature and, therefore, straightforwardly fall in the grouping of productive instruments. In these cases, the apparatus is inextricably blended with the world, and the phenomena are attributes of the apparatus/world complexes. Examples of this subclass are the apparatus that Davy used to isolate sodium in the metallic state via electrolysis, Faraday's electromotor, Ohm's thermocouple, and particle accelerators like the LHC at CERN, when they are used to produce phenomena not known in nature.

b. *Instruments as tools for purifying, controlling and/or manipulating phenomena*: this subclass fits well in the grouping of modelling instruments, because they are used to control, re(produce), and manipulate phenomena. The natural sciences offer a lot of examples, including Roentgen's use of his experimental set-up to influence the X-rays and make them behave in a certain way,[136] Atwood's machine – an instrument to study falling bodies in an artificial setting – and the use in artificial settings of certain enzymes as catalysts for certain chemical reactions. Tsou (2012) emphasises how pharmacological drugs are often used as tools to manipulate neurobiological phenomena in psychopathology.

c. *Model organisms*: this subclass fits the modelling grouping, but not entirely. In fact, model organisms are different from material models because they are representing a phenomenon but, at the same time, they are a representative part of the phenomenon. In other words, they are both samples of the phenomenon under investigation and representational artefacts that have undergone a process of standardisation. In this sense, they embody certain selected features that are materially relevant to experimental manipulation and, therefore, determine partly the outcome of the epistemic interaction (Leonelli 2007). From the point of view of their function, it might be said that they can sometimes have a productive role, in those cases in which they give rise to phenomena that do not happen in the wild (for example, the loss of reproductive power of some species in captivity).

d. *Material models:* this subclass fits in the modelling grouping. Usually, the phenomenon reproduced is not the same as the natural one, since material models are distinctively representational artefacts aimed at representing a phenomenon in an artificial setting to investigate its properties. Still, these models are part of the experimental process, since they simulate an interaction between epistemic subject, phenomenon, and environment. Studies on material models, such as Watson and Crick's double helix model, highlight how the material features of the models crucially determine the possibilities of manipulation and interaction between epistemic subject and physical system (Charbonneau 2013, Giere 2010, Griesemer 1990).

[136] Cf. Heidelberger (2003: 147). Heidelberger makes a difference between Roentgen's productive and constructive use of his experimental apparatus.

e. *Detecting tools:* these instruments, such as the litmus paper detecting the presence of acidity, fall quite clearly in the quantifying grouping broadly understood, since they interact with the phenomenon to detect the presence of a certain property, which can be considered as a minimal form of quantification.

f. *Measuring instruments*: these instruments, such as thermometers and clocks, are paradigmatic examples of quantifying instruments, since they represent one or more properties of the phenomenon by means of a physical interaction taking place within the instrument, which often, but not always, must physically interact with the phenomenon of interest.

Once again, I emphasise that these classifications should not be taken as exhaustive or final.[137] Nor my continuum should be regarded as a rigid frame, or one that accounts for all aspects of materiality in experimentation.[138] Yet, in the next section, this continuum will provide one important dimension for the analysis of those cases in which scientific instruments seem to have a constitutive character.

4.3.2 Material constitutivity as a trade-off

In the previous section, I categorised scientific instruments according to both their epistemic function and some of their structural features. I ordered these groupings on a continuum depending on the extent to which the materiality of the instruments, i.e., certain material features they have, is relevant to the epistemic outcome of experimentation, that is, to the product of the interaction between instrument, epistemic subject(s), and environment. More precisely, this continuum represents different ways in which the materiality of scientific instruments constrains experimental manipulations, thus constituting another source of justification to the representational character of experimentation. In other words, certain material features of the apparatus, in some circumstances, constrain the experimental manipulation so that the way in which the epistemic outcome of that manipulation is represented is partly justified by those material features, rather than by features of the phenomenon under inquiry.

This material source of justification is distinct from (although, in some cases, correlated with) the degree of theory-dependence of an experimental manipulation, as I discussed it in chapter 1 (section 1.3.3). At that point, I suggested that all experimentation has a representational character, i.e., all

[137] For instance, both certain detecting tools and some measuring instruments can have, to some extent, productive effects, as in the classic problem of measurement in quantum mechanics.

[138] For instance, the impact of standardisation of the material features of all groupings and subclasses of instruments on the epistemic outcome of the experimental interaction has been largely underemphasised here. My only justification for this omission, pragmatic reasons aside, is that standardisation might prove equally relevant, in this context, for all the classifications spelled out.

experimental outcomes are, in some sense, representations. I also suggested that the extent to which experimentation has such a representational character is a matter of degree and depends on the relationship that different varieties of experimentation have with theory. Based on this assumption, I introduced a conceptual space in the form of a continuum, on which I placed different types of experimentation according to the degree of theory-dependence of their representational character. I present again this continuum in fig. 4.4.

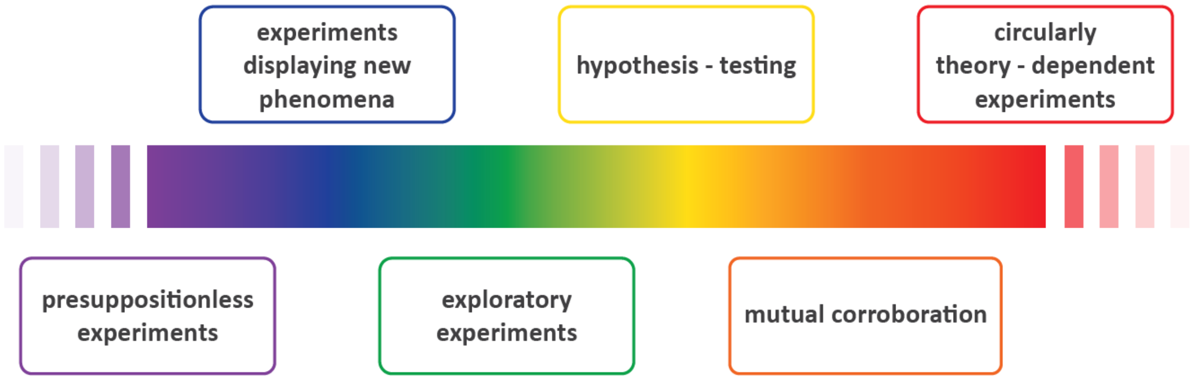


fig. 4.4: continuum of theory-dependence of the representational character of experimentation. Every item represents a different sort of experimentation according to the extent to which theory justifies the representational character of the experimental set-up

The continuum of materiality in fig 4.3 does not merely represent the specular version of the continuum of theory-dependence in fig 4.4. These two dimensions are neither mutually exclusive, nor do they express features that are distributed according to inverse proportions. For instance, productive instruments – at the left hand-side of the materiality continuum – are not always closer to theory-freedom – i.e., towards presuppositionless experiments, at the left hand-side of the theory-dependence continuum – as one might expect by simply juxtaposing the two continua. Examples of phenomena produced by means of apparatus/world complexes, such as the LHC at CERN, often show that – despite the space of material possibilities that the instrument opens by generating new phenomena (Mormann 2012) – the experimental manipulation presupposes thorough theoretical justification, especially in cases in which instrumentation gets very complex (Radder 2003), or when instruments are used to the limits of their capacity (Chalmers 2003). On the other hand, typically quantifying instruments, such as measuring instruments, can sometimes be devised even when some deeper theoretical knowledge justifying their workings is lacking (Chang 2004). This becomes evident, for instance, in those cases in which measuring instruments, are used in combination with productive instruments, as in the case of Ohm's experiments with the thermocouple.

Ohm's experiments with the thermocouple were certainly performed on the basis of several theoretical commitments, including Coulomb's torsion law and Ampére's force law, and aimed at testing the hypothesis of a certain relationship between intensity of the electric stream and resistance. This hypothesis was not even formulated for the first time by Ohm, since it had already been subject to scrutiny for a long time.[139] However, the fact that he relied on a novel form of measurement coordination between the quantity of exciting force and his measurement procedure based on the apparatus torsion balance+thermocouple provides a reason to place his experimental practice between the categories of hypothesis-testing and exploratory experiments, rather than fully in the category of hypothesis-testing (fig. 4.5). Indeed, one meaning of 'exploratory experiments' is that of 'domesticating' phenomena, that is, reproducing natural phenomena in an artificial setting so that they present purified and simplified features (Harré 2003). In this sense, Ohm domesticated the electric stream he used for his hypothesis-testing by using an instrument, the thermocouple, whose workings were rather poorly understood from the theoretical point of view.

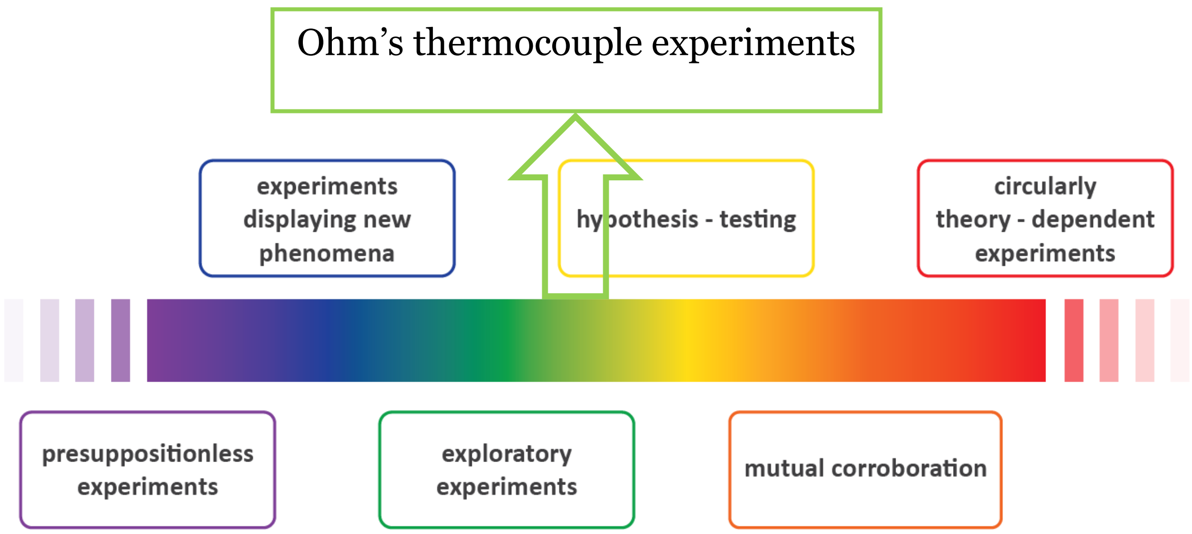


fig. 4.5: The place of Ohm's thermocouple experiments in the continuum of theory-dependence

The thermocouple was a relatively new instrument and, as we saw in section 4.2.3, the interpretation of its thermomagnetic effects in terms of electric effects was not a settled issue at the time Ohm used this instrument in his experimental practice. Yet, as I emphasised, Ohm spent limited time in testing the reliability of the thermocouple in producing electric streams of intensities proportional to different temperature differentials. It may be said that, at least in part, Ohm relied on the 'proper functioning' of the thermocouple in producing a stable electric stream, against the background of the justification

[139] Cf. above, section 4.2.1.2.

provided by his more reliable measurement instrument, i.e., the torsion balance. Both instruments were crucial to the outcome of Ohm's experiments, the balance as a measurement instrument and the thermocouple as a productive instrument. The productive use of the thermocouple to investigate the effects of resistance on the electric stream presupposed the understanding of a newly produced phenomenon, i.e., the thermomagnetic effects of differences in temperature, in terms of another phenomenon, i.e., the electric current on which Ohm was experimenting. In other words, the thermocouple was an apparatus/world complex, since it generated a phenomenon, i.e., a thermoelectric current, that could not be found as such in nature and was poorly theoretically understood (fig. 4.6).

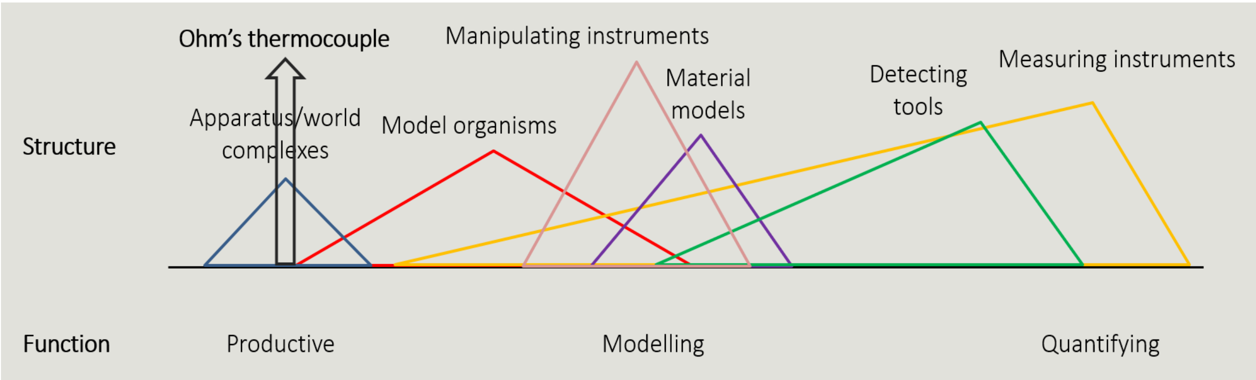


fig. 4.6: The place of the thermocouple used by Ohm in the continuum of materiality

The two dimensions I introduced, that is, the continuum of materiality (x) and the continuum of theory-dependence (y), can be represented as the two axes of a Cartesian space (fig. 4.7). Both axes have a direction, indicating the extent to which an instrument in a specific experimental context relates to these two dimensions. In this space, I placed Ohm's thermocouple used as a productive instrument in the context of his experiments on conductivity and resistance, and some other examples among those I mentioned in this section. In the light of my analysis, the thermocouple is located in the upper-right quadrant, that is, the area in which the materiality of instrumentation has a significant role in justifying features of the epistemic product of experimentation while, at the same time, the theoretical background does not do the heavy-lifting of justifying the representational character of the experimental setting, where the experimental setting, in this case, includes the use of the thermocouple for productive purposes.

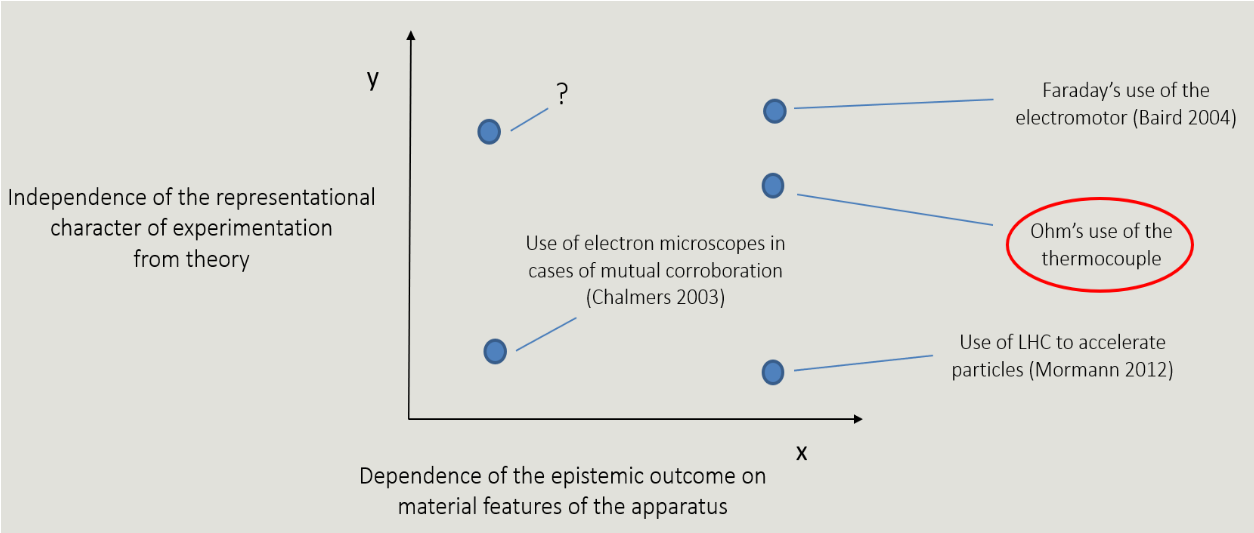


fig. 4.7: Ohm's thermocouple as a productive instrument in the Cartesian space representing both dimensions of materiality (x) and theory-dependence (y)

Although I am using this Cartesian space to analyse only the case of Ohm's thermocouple, I think it is useful to provide an initial understanding of how the material base of scientific inquiry and, more specifically, scientific instrumentation, can have a constitutive character in coordination. This character can be conceptualised as a trade-off between the extent to which the representational character of an experimental set-up is justified by theoretical background, and the extent to which material features of instrumentation constrain the experimental manipulation and, therefore, justify some feature of the epistemic product of experimentation, as shown in fig. 4.7. In other words, the more the materiality of instrumentation determines features of the experimental outcome and the more the justification of the representational character of experimentation is independent from theory, the more the material apparatus is constitutive in a form of coordination. In this way, we can express the fact that some instruments provide certain conditions to identify the referential relationship between certain epistemic objects (e.g., measurement outcomes, empirical regularities, etc.) and how the epistemic subject(s) represent them in some situated contexts of scientific inquiry. Their constitutive character is of a different kind than the theoretical components analysed in this and previous chapters, since it does not come from constraining a space of theoretical possibilities, but from delimiting a material space of intervention by enabling certain physical interactions between apparatus, epistemic subject, and environment, and not others.

4.3.3 The material constitutivity of Ohm's thermocouple

The constitutive character of Ohm's thermocouple in measurement coordination cannot be considered in isolation, but together with the rest of the measurement procedure and in the specific context of Ohm's experimental practice. In this sense, the proper functioning of the thermocouple in producing a certain phenomenon, Coulomb's law as the measurement law underlying the workings of torsion balance, and Ampère's force law as justifying the inference from readings of the torsion balance to outcomes of exciting force, were involved in justifying the measurement coordination between values of exciting force and the procedure to measure them.

In terms of the features of constitutivity, all three components had a relatively high quasi-axiomaticity (QA), since their epistemic function was kept fixed in Ohm's inquiry, even if they are very different kinds of components, two theoretical and one material. The same holds for their relatively high generative potential (GP), considering the equal importance of their justificatory function for Ohm's measurement coordination, since they underpinned the same amount of conceptual dependence-relations supported by the concept of exciting force in Ohm's inquiry. What about their empirical shielding (ES)? The meaning of 'empirical shielding' cannot be the same when it is used to describe a property of theoretical components, and when it is used to describe a property of a material component. What does it mean for an instrument to be empirically shielded? An instrument cannot express any conceptual content with any degree of generality, nor is it in any way subject to abstraction, although it can contribute to the abstraction and generalisation of a phenomenon, by providing certain constraints for its production, modelling, or quantification. The only plausible understanding of empirical shielding with respect to a scientific instrument is in terms of the extent to which it is possible to directly test and assess the proper functioning of the instrument, that is, its capacity of working effectively for the epistemic goal it is designed for. If the proper functioning cannot be assessed, the instrument has maximal empirical shielding. For instance, in cases of measuring instruments whose workings are poorly or not at all theoretically understood, it is hard to assess whether the instrument is properly working towards the epistemic goal of quantification. On the other hand, in the case of the thermocouple – a productive instrument – its proper functioning could be assessed relatively easily, to the measure that it could reproduce a certain phenomenon (i.e., a stable electric stream with intensity roughly proportional to temperature differential). Therefore, the empirical shielding of the thermocouple is slightly lower than the other two which, albeit being empirical regularities, were shielded, to some degree, from being directly testable within the context of Ohm's experimentation (fig. 4.8).

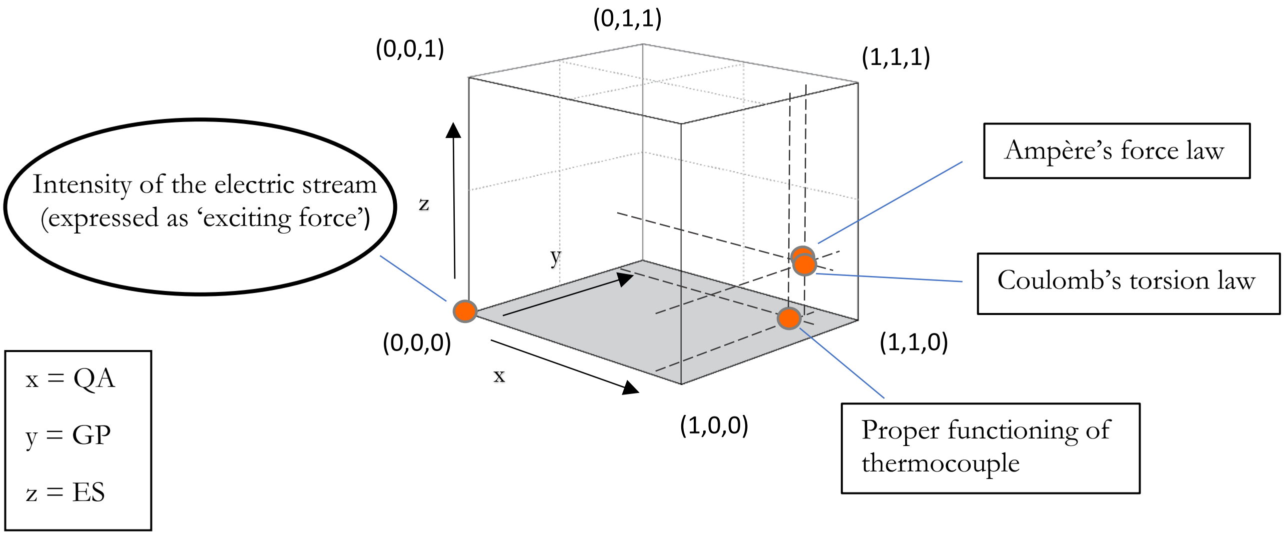


fig. 4.8: Graphic representation of the constitutivity of the thermocouple, of Coulomb's torsion law, and of Ampère's force law in the measurement coordination between the intensity of the electric stream and its expression by means of the quantity of exciting force. As usual, the three features are quasi-axiomaticity (QA), generative potential (GP), and empirical shielding (ES)

If we zoom out to from the relationship between the components involved in measurement coordination and try to represent how they relate to general coordination, we see that measurement coordination is itself a component of the general coordination (fig. 4.9). Measurement coordination has, in fact, a high quasi-axiomaticity and generative potential, given that Ohm held fixed the coordination between exciting force and his measurement procedure based on the thermocouple and the torsion balance to establish his law, and the law itself is conceptually supported by this measurement coordination. The same holds for the theoretical principle by which tension is equally distributed in a circuit, which has, however, a higher degree of empirical shielding than measurement coordination, since the former is shielded due to its elevation from inductive generalisation to deductive axiom, and the latter comprises epistemic components that, as we have seen, do not have high degrees of empirical shielding.

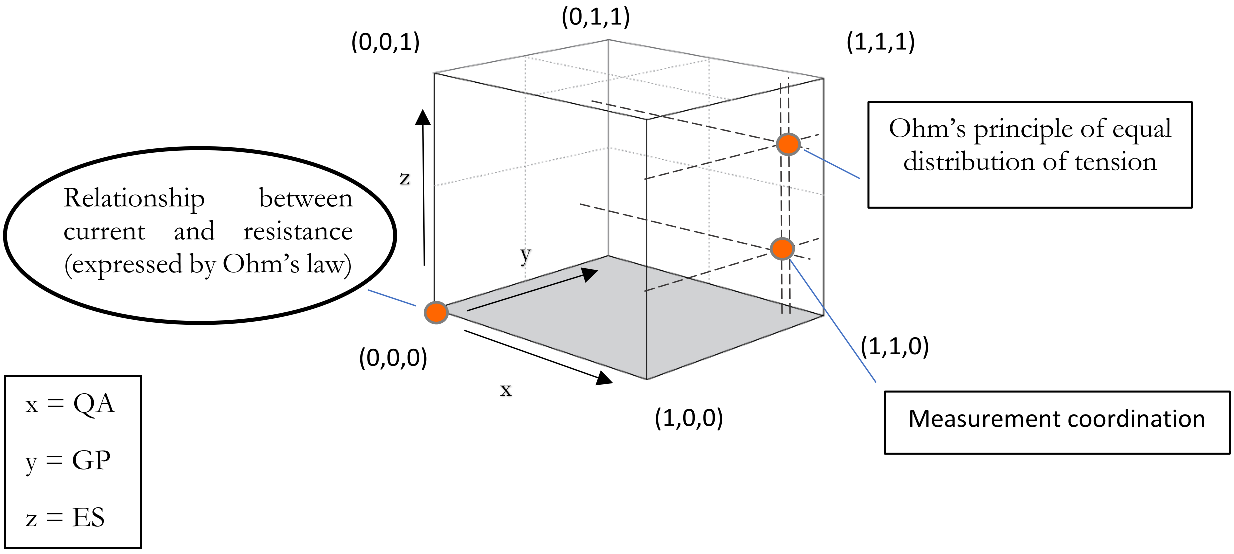


fig. 4.9: Graphic representation of the constitutivity of measurement coordination and of the concept of tension in the general coordination of the empirical relationship between intensity of current and resistance with Ohm's formulation of this regularity

4.3.4 Final note: the role of approximation in calibration and measurement

In this chapter, I have suggested that in some epistemic contexts, scientific instruments with specific epistemic functions can be constitutive within a coordination process. Their constitutivity is of a material kind, in contrast with the theoretical principles analysed in previous chapters. More specifically, the material constitutivity can be conceived as a trade-off between the theory-independence of the representational character of the experimental context and the dependence of the epistemic outcome of experimentation from material features of the apparatus. In the specific case of Ohm's measurement practice, I showed how the thermocouple had, to some degree, a constitutive character in the measurement coordination on which Ohm relied. To emphasise its material constitutivity, I analysed the role of the thermocouple in comparison to both the role of the theoretical assumption of equal distribution of tension in the general coordination between Ohm's law and the regularity it expresses, and the role of Coulomb's torsion law and Ampère's force law for the achievement of measurement coordination. To conclude, I would like to briefly gesture towards a further epistemic component involved in his scientific inquiry: the principle of approximation. This component seems to have had a crucial role both in Ohm's calibration and in the assessment of the proper functioning of the thermocouple.

In section 4.1.4, I presented calibration as a complex modelling activity that makes use of several theoretical, statistical, and pragmatic assumptions to empirically account for uncertainties and systematic errors of the measurement process. In general, the great majority of the inferential assumptions involved in calibration seem to be local, as in the case of measurement laws, since they refer to the specific interaction taking place in the measurement procedure under consideration. However, a more general sort of assumption may appear to be required by the process of calibration as a modelling activity in general. Van Fraassen (2008: 52) suggests that the use of idealisations in science necessarily requires assuming what he calls the 'principle of approximation', which he defines as follows: "If certain conditions follow from the ideal case, then approximately those conditions will follow from an approximation to the ideal case".[140] Here, I just want to show how this principle connects to the activity of calibration. In a nutshell, calibration as the activity of modelling the measurement process involves a combination of more basic activities. Central to calibration are the activities of idealising and de-idealising from certain experimental, statistical, or theoretical assumptions, often expressed in mathematical form. These activities are required, so to speak, to 'translate' the constraints or adjustments prescribed by these assumptions into physical interventions on the measurement procedure, so as to minimise the role of systematic measurement errors and uncertainties and provide stronger justification to the inferential step from instrument reading to measurement outcomes. The principle of approximation seems to express a general precondition that we must assume to be able to reason in terms of idealisations and de-idealisations. I will explore this interpretation of the principle in the next chapter.

[140] Van Fraassen attributes the first formulation of this principle to Reichenbach (1956: 93-95).

CHAPTER 5

5.1 Counterfactual reasoning, coordination, and the Hardy-Weinberg principle

5.1.1 Constitutive elements do not have to be domain-specific

In the last chapter, I showed that certain epistemic components of a scientific framework can be, to some degree, constitutive of a form of coordination without being theoretical in character. More specifically, I argued that scientific instruments in specific circumstances can exhibit some of the typical features of constitutive elements. Although the extent to which they are constitutive can be expressed by the same features, these material components differ from the theoretical ones in that they do not constrain scientific inquiry by delimiting a space of conceptual possibilities, but by delimiting a concrete space of possible physical interactions. Yet, both kinds of elements share the trait of domain-specificity, that is, the fact that their constitutive character is limited to a scientific framework and, especially in the case of material components like specific scientific instruments, to situated scientific inquiries within which a form of coordination is established.

In this chapter, I will investigate whether some epistemic components can have a constitutive character without being necessarily tied to a specific scientific framework or empirical domain. I will suggest that some components are involved in the coordination between abstract representations and phenomena in a rather different fashion than the domain-specific components that I have analysed so far. Yet, my argument relies on a case study which *prima facie* seems to have many traits in common with the theoretical principles analysed in previous chapters. I will investigate the history and epistemic function of the Hardy-Weinberg principle (HWP), which is central to the workings of population genetics. Even though some historians and philosophers of science have interpreted it, in the light of its history, as a component of an experimental system (Rheinberger 2013, Rheinberger & Müller-Wille 2017), the HWP is usually regarded by both scientists and philosophers as a fundamental principle of the theory of population genetics (Gillespie 1998, Sober 1984). The polarity between practice-based scientific frameworks or units (e.g., experimental systems) and theory-based ones (i.e., theories) does not really matter for my purposes in this chapter. My analysis of the HWP is mainly focused on understanding the role of constitutive components that do not seem to be framework-relative but, rather, domain-independent. Analysing these components in abstraction from any specific case study would be an arduous task, because their role can be much more easily scrutinised in the light of their relationship with a domain-relative element.

In the rest of this section, I introduce the topic of the domain-independence of certain epistemic elements. More specifically, I focus on counterfactual reasoning as an example of a general reasoning ability and on how its cognitive and epistemic role has been conceptualised in some philosophical and scientific debates. In addition, I discuss the relevance of the HWP to illuminate another side of the issue of coordination. In section 5.2, I provide an analysis of the epistemic history and role of the HWP in population genetics. I deploy the three features of quasi-axiomaticity, generative potential, and empirical shielding to show to what degree the HWP can be viewed as constitutive in population genetics. What I suggest is that it is involved in a form of coordination in that it represents a counterfactual ideal state of equilibrium that enables the representation of concrete phenomena, in this case, changes in patterns of genetic distributions, as deviations from expected frequencies due to evolutionary causes. However, the HWP is itself constituted, since its formulation and use are a product of counterfactual reasoning, which, in turn, requires the assumption of an idea of stability. In addition to that, the use in practice of the HWP requires to de-idealise from its strong idealisations, thus implying the assumption of the principle of approximation. In section 5.3, I outline the epistemic function of the principle of approximation and of the idea of stability and argue that they are qualitatively different from domain-specific principles. Finally, I suggest that they are domain-independent preconditions underlying different reasoning abilities aimed at representing phenomena across scientific frameworks.

5.1.2 Counterfactual reasoning as a tool for scientific inquiry

When it comes to analysing certain epistemic practices rather than individual theories, it seems that some general assumptions underlie certain reasoning abilities that scientists use across disciplines and historical periods. In this chapter, my goal is to understand whether these components provide epistemic justification for coordination in a different fashion than domain-specific theoretical presuppositions. In other words, my question is whether there are epistemic components that are constitutive not because they provide fundamental conceptual resources in a theory or domain, but because they detain a normative force across domains by regulating and providing justification to general epistemic practices while, at the same time, contributing to the coordination between phenomena and their representation. Certainly, scientific practice and theorising involve forms of reasoning as diverse as, for instance, causal reasoning, statistical reasoning, experimental thinking, or classificatory thinking. On the one hand, historians and philosophers of science have traced the historical emergence of certain general 'styles' of scientific reasoning based on different strategies of knowledge acquisition, justification criteria, and objectivity standards (e.g., Crombie 1994, Hacking

1992). On the other, cognitive-oriented analyses of science have focused on the dynamics by which different cognitive faculties work and interweave, for instance, in the attempt to understand the role of faculties such as imagination in the process of concept-formation in the sciences (e.g., Nersessian 1992, 2010).

Counterfactual reasoning is one of the fundamental reasoning abilities widely used in the sciences. Ever since the 1980s, psychological literature has provided evidence that normative considerations influence how easily people have access to various alternatives for consideration in counterfactual reasoning, thus highlighting the role of normative judgments in regulating the accessibility of counterfactual scenarios (Kahneman & Tversky 1982, Kahneman & Miller 1986). The ability of imagining what something would be like, were certain conditions to hold, is almost trivially required by epistemic activities such as formulating hypotheses or making predictions. In addition, there are rich philosophical traditions that focused on the role of counterfactual reasoning, respectively, in thought experimentation and its connection with real experimentation, and in scientific modelling.[141]

As I mentioned in chapter 1 (section 1.2.3.2), Buzzoni (2008, 2013a, 2013b, 2018) interprets scientific thought experiments as counterfactual anticipations of the experimental apparatuses in terms of their conceptual content, where the understanding of the causal connections is made possible by using counterfactual conditions. In thought experiments we imagine ideal entities that provide rules, norms, criteria without which "we would be unable even to *conceive of* the imperfections and empirical deviations typical of *real* situations" (Buzzoni 2013a: 101).[142] Therefore, counterfactual reasoning in general is to be regarded as a reasoning ability that is not only necessary for the activity of thought experimentation, but also for actual experimentation, given that all real experiments can be thought of as realisations of thought experiments. Stuart (2017) builds on Buzzoni's view and argues that thought experiments can fix the meaning of different theoretical structures (laws, equations, models, concepts, etc.) by using imaginary examples. Stuart further characterises Buzzoni's appeal to counterfactual reasoning as a transcendental precondition for thought and real experimentation by invoking imagination as a mental faculty that involves counterfactual reasoning as its constitutive component.[143]

141 Especially the literature on causal modelling focuses extensively on the role of counterfactual reasoning. For a recent overview, see Beebee et al. (2017). For a discussion of the role of counterfactuals in causal reasoning in the biological sciences, see Weber (2017).
142 Emphasis in the original.
143 In this sense, Stuart's view is in resonance, on the one hand, with Nersessian's contribution focusing on the synthetic role of imagination with respect to empirical representations, and, on the other, with both

Counterfactual reasoning is also crucial to a variety of modelling practices in the sciences.[144] Not only is it typically required in causal modelling, but it is also involved in the process of surrogative reasoning underlying the practice of representational modelling in general, which is underpinned by inferences of the form 'if the model *x* represents the system *y* for the purpose *z*, then the property *A* of the model corresponds to property *B* of the system according to interpretation *ϱ* of the representation relation' (Contessa 2007).

In this chapter, I cannot do justice to any of these bodies of literature. Rather, I will only consider and develop some aspects they emphasise, with respect to the rich role that counterfactual reasoning plays in scientific inquiry and to the distinctive traits of domain-independent reasoning abilities required for performing certain epistemic practices. These aspects will be central to my analysis of the Hardy-Weinberg principle and of the components which underpin its formulation and use in population genetics.

5.1.3 The Hardy-Weinberg principle: epistemological perspectives

The HWP is a theoretical tool deployed in a wide range of modelling practices implemented by population geneticists (Gillespie 1998, Hartl & Clark 2007, Russell 2010). Its simplest formulation assumes the presence of only two variants of the same allele, dominant and recessive. It states that, in the absence of evolutionary forces (natural selection, mutation, migration), in an infinite population of sexually reproducing diploid organisms that mate randomly, the genotypic frequencies of the offspring generation will be distributed according to the following frequencies:

$$AA\colon p^2 \qquad Aa\colon 2pq \qquad aa\colon q^2$$

where *A* and *a* are the two allelic variants which determine the three possible genotypes (*AA*, *Aa*, and *aa*), whereas *p* and *q* are the allele frequencies in the parent generation.

With the establishment of the Modern Synthesis in the 1930s,[145] the use of the HWP became customary as a description of the *equilibrium state* of (at least a class of) genetic populations, and it is often interpreted as expressing the conditions under which no external *force* acts on a system (Stephens 2004, Hartl & Clark 2007). Given this role, genetics textbooks and papers generally characterise the HWP as a *zero-force model*, and analogies between this role and that of the first law of motion in

Lichtenberg's and Kuhn's conceptual constructivism, which regarded thought experimentation as a way to access both our concepts and the world. Cf. Fehige & Stuart (2014).

144 Here I refer to models in the most general way possible, as to include everything that is regarded or used as a representation by a model user (cf. Teller 2001).

145 See above, chapter 3, section 3.1.2.

Newtonian mechanics are frequent (Mayo 2008, Ruse 1971, Sober 1984). Yet, sceptical attitudes towards the HWP as a zero-force state, and towards the understanding of evolution as a theory of *forces* in general, have emerged (Earnshaw 2015; Matthen & Ariew 2002, 2009; Walsh 2007). In addition, some researchers have emphasised how certain assumptions of the HWP are problematic for its role of zero-force state (Li 1988; Stark 2006a, 2006b, 2007; Stark & Seneta 2013). Consequently, options alternative to the HWP have been investigated, with the aim of overcoming its limitations, by either reconceptualising the notion of equilibrium state in a dynamical sense (Bosco et al. 2012) or providing a characterisation of the zero-force condition for all evolutionary systems (McShea & Brandon 2010).

In the following sections, I take cues from the history of the role of the HWP in population genetics and from the current prospects of replacing it, in order to better understand its epistemic function. My goal is to show that certain epistemic components are constitutive in a qualitatively different way from the domain-specific elements analysed in previous chapters. Indeed, it has already been argued that the mathematical ratios expressed by the HWP do not – in and of themselves – justify any inferences to phenomena, but, at most, they constrain which interpretations of the formalism are possible given the variables and their interaction (Millstein et al. 2009). Therefore, it is not specifically in virtue of some mathematical concepts or principles that the coordination between frequency changes in genetic populations as a concrete phenomenon and their abstract representation in mathematical form as deviations from expected frequencies can be established. Rather, I will highlight that it is the counterfactual character of the HWP that comes to work when the formalism is used to interpret a frequency distribution and to make hypotheses on the deviations from expected frequencies. In other words, I will argue that the counterfactual character of the HWP justifies the coordination between concrete changes in genetic frequency distributions and their representation in mathematical form as deviations from expected frequencies. In addition, I will show that the possibility to make inferences from these abstract representations to phenomena by using the HWP relies on two domain-independent assumptions, i.e., approximation, required to de-idealise from the highly idealised HWP, and stability, which is required to model a phenomenon via counterfactual reasoning. I will suggest that these assumptions are best understood as domain-independent principles underlying the reasoning abilities that epistemic subjects deploy to idealise and model.[146] Yet, I will

[146] My interpretation here is close to what Chang (2008, 2009) suggests. It might be pointed out that Chang's own transcendental principles (cf. above, chapter 1, section 1.2.3.3) can themselves be considered as instantiations of counterfactual reasoning in domain-independent assumptions relative to different epistemic activities. For instance, the principle by which it is required that something is discrete in order to count it

show that the HWP and the assumptions of approximation and stability can be fruitfully analysed in terms of the features of quasi-axiomaticity, generative potential, and empirical shielding.

Before turning to my own analysis, I must point out that a prominent epistemological understanding of the HWP characterises it as an *a priori* element of population genetics, when suitably formulated (Sober 1984). My argument that the HWP exhibits, to some degree, a constitutive character is mainly aimed at showing the explanatory potential of my view of constitutive elements in science. Therefore, my original contribution in this chapter is not aimed at having a direct impact on philosophical perspectives in population genetics. However, as it happens, I will endorse specific perspectives and implicitly or explicitly reject some others, with reference to debates in biology and philosophy of biology. In the light of this consideration, on the one hand, and to the aim of bridging the debate on constitutive principles and those in the epistemology of the biological sciences, on the other, in the next section I will also provide grounds to claim that my understanding of the HWP as constitutive is not in principle incompatible with Sober's view of the HWP.

5.2 The epistemic role of the HWP in the history of population genetics

5.2.1 <u>The HWP as a case study</u>

As I highlighted above, the HWP is often viewed as a zero-force model expressing the equilibrium state of genetic populations. In actual practice the HWP – and its more specific variants – provides the proportions to which allele frequencies in real populations (obtained, for instance, via 'gene counting') are tested against by means of statistical methods, then allowing to formulate hypotheses to explain the extent and reason of any departure from Hardy-Weinberg proportions. It is for this reason that the HWP is often addressed as a zero-force model that describes the equilibrium state of a genetic pool holding under certain conditions. By adding further conditions to the general formulation of the HWP, i.e., by introducing parameters through which, for example, rate and frequency of mutation in a real population can be modelled, more realistic models can be developed.

I have also mentioned that, according to a classic philosophical interpretation, the HWP, when suitably formulated, is *a priori* (Sober 1984). According to Sober, the character of zero-force model of the HWP entails that it is devoid of empirical content and, thus, that it cannot be empirically refuted. This claim supports the view that there are laws in biology, even if they are not empirical laws (Elgin 2003). In

expresses the impossibility of conceiving a transgression of the norm imposed by the principle while performing a certain activity, i.e., of imagining a counterfactual scenario in which that condition does not hold.

this section, I am not contributing to the debate on the nature of biological laws, nor to the one on the legitimacy of an empirical requirement for scientific laws. Rather, my aim is that of assessing the epistemic status and function of the HWP in the historically situated development of population genetics, and according to its role in the current workings of population genetics. What I will show is that my account based on quasi-axiomaticity, generative potential, and empirical shielding allows me to capture the extent to which the HWP is constitutive in the form of coordination in which it is involved. As for the principles of natural selection in chapter 3, I will suggest that the HWP has a certain degree of constitutivity, even though its specific epistemic role is different from the one Friedman focuses on, since the HWP does not constitute any 'first-order' empirical law. Nor – I will suggest – does it count as constitutive merely because it can be viewed as *a priori* in Sober's sense.

5.2.2 Origins of the HWP

Darwin's theory of evolution by natural selection lacked a specific mechanism of heredity that could explain the transmission of characteristics from parent generation to offspring. Darwin thought that the parents' characters 'blended' in the offspring generation, but this blending required a much greater amount of variation for natural selection to operate than he had previously hypothesised. Only in 1868 did Darwin suggest the 'hypothesis of pangenesis', according to which each part of an organism produces small 'gemmules' transmitted to the offspring via reproduction. Darwin's account was based on the importance of small individual differences as the main source of the variations on which natural selection could operate. This gradual view of evolution by natural selection found its early opponents in Huxley and Galton, who argued that it was mainly discontinuous variations (or 'sports') that could be acted upon by natural selection. Consequently, they thought that evolution would proceed by discontinuous leaps, or 'saltations'. This was the initial phase of the debate opposing gradual and discontinuous evolution, which terminated only with the consolidation of population genetics and the synthesis of Darwinism and Mendelism.[147]

The distinction between continuous and discontinuous variations and its relation to evolution radicalised with the next generation of scientists and turned into the famous conflict between the biometricians – Pearson and Weldon in particular – and Bateson. The latter strongly opposed the biometricians on all their central tenets: the importance of individual differences, the gradualism of evolution, the fundamental role of statistical techniques, the neglect of the physical mechanism underlying inheritance. The contrast was already very sharp when, in 1900, De Vries, Correns, and

[147] Cf. above, chapter 3, sections 3.1.2 and 3.2.2 (especially footnote 77).

von Tschermak rediscovered Mendel's work on inheritance in peas. The different interpretations of Mendel's work fuelled the debate. Bateson's work on Mendel's theory and De Vries' mutation theory led to the association of Mendelism – which was based on particulate units of inheritance, and thus seemed to imply discontinuous variations – with evolution 'by jumps', as opposed to Darwinian continuous variations and gradual evolution.

Well before these events, Mendel had investigated the consequences of self-fertilisation in pea plants and in 1865 he concluded that hybrid parents with respect to a specific character would produce offspring exhibiting that character according to specific proportions. He observed that, in the offspring generation, the pure forms *A* and *a* and the hybrid form *Aa* were distributed according to the proportion *1A:2Aa:1a* and postulated the so-called 'factors' to explain the combination of traits in reproductive cells. Consequently, he deduced that there must be two factors in a combination, which segregate (i.e., separate) in the production of gametes, and that the factors do not blend, but remain distinct and do not influence each other in segregation. In addition to these two assumptions (law of segregation and law of independent assortment), he also assumed that of the two pure forms one is dominant and the other recessive with respect to the phenotypic manifestation of the trait (law of dominance). Although both the law of dominance and that of independent assortment were later found untenable, Mendel's great heritage is that he, for the first time, "had hypothesized that inheritance of a trait was particulate, that the units of inheritance did not change from generation to generation, that they were contributed equally by an organism's two parents, and that each parent contributed its unit at random from the two it contained" (Mayo 2008: 249). Yet, after its rediscovery, Mendelism remained controversial for many years (Olby 1966).

In the midst of the controversy between biometricians and Mendelians, the formulation of the HWP came about as a rather simple mathematical derivation from the Mendelian scheme (Diaconis 2002, Edwards 2008).[148] The zoologist and geneticist R. C. Punnett asked his Cambridge colleague, the mathematician G. H. Hardy, whether there was any intrinsic reason to expect a dominant character to continually increase in frequency over time, to the detriment of the recessive.[149] Hardy (1908)

148 See also Provine (1971: 131-136).

149 Punnett made this request to Hardy after his lecture on "Mendelism in relation to disease" to the Royal Society of Medicine in 1908 in London. During that lecture, G. U. Yule pointed out that, if brachydactyly is a dominant character in man, by assuming random mating we should expect a proportion of 3:1 brachydactyly : normal, which was not what the evidence showed. Yule seemed to assume, as he did already in his 1902 paper, that the frequency of the two characters *A* and *a* in a population was of one half, inevitably leading to the 3:1 proportion in the offspring generation. Ironically, by demonstrating that under the condition of random mating and with the assumption of complete dominance, the offspring generation of hybrid parents would preserve the proportions of 3:1 dominants : recessives, Yule had himself reached the

demonstrated that, assuming a large population of randomly mating individuals, if the parental genotypic proportions for the three genotypes *AA*, *Aa*, *aa* were p:2q:r respectively, then they would be $(p + q)^2$: $2(p + q)(q + r)$: $(q + r)^2$ among the offspring. He highlighted that the condition for the distribution to remain the same from parent to offspring generation was that $q^2=pr$, and that, since this relationship between q, p, and r remains stable independently of the values assumed by p and r, the distribution will be stable after the second generation, provided no changes in the conditions occur. In Hardy's own words, "there is not the slightest foundation for the idea that a dominant character should show a tendency to spread over a whole population, or that a recessive should tend to die out" (Hardy 1908: 49). The German physician Wilhelm Weinberg independently obtained the same result in the same year.[150]

5.2.3 A cognitive-historical analysis of the epistemic function of the HWP

5.2.3.1 Quasi-axiomaticity

Even though neither Hardy nor Weinberg regarded the formulation of the principle as a significant contribution (Crow 1999), and in the beginning its importance was not unanimously acknowledged,[151] the HWP quickly gained importance in the work of the 'founding fathers' of population genetics: R. A. Fisher, S. Wright and J. B. S. Haldane. Despite it took some time for Mendelism to become uncontroversial and the debate between Mendelians and biometricians prolonged well into the 1920s, the use of Hardy-Weinberg proportions as equilibrium principle was clearly a crucial feature of the early work of those scientists who conjugated Mendelism with Darwinism and with the statistical 'style' introduced by the biometricians. In what follows, I will briefly rehearse some well-known pieces of historical evidence, in support of the claim that the HWP played a fundamental epistemic role during the phase in which population genetics was brought to full life.

first, albeit restricted, formulation of the Hardy-Weinberg principle (Edwards 2008). However, he did not examine this result any further in that 1902 paper, nor the assumption of random mating that he himself introduced.

150 Weinberg's work went largely unnoticed until 1943. For an appreciation of Weinberg's contributions, see Crow (1999). Edwards (2008) highlights that Pearson (1904) is sometimes mentioned as the first paper that showed the Hardy-Weinberg proportions for the special case of gene frequency of one half. In 1909, eventually Pearson explicitly derived the HWP and referred it to the work of Hardy. Castle (1903) is also often indicated as a precursor of the Hardy-Weinberg principle because, as Provine (1971: 133) points out, he expressed in non-mathematical terms the equivalent of the equilibrium principle for a single locus with two alleles. For an alternative view on Castle's contribution, see Edwards (2008: 1149).

151 For instance, in Bateson's (1909) account of Mendelism there is no recognition of Hardy's law (cf. Edwards 2008: 1147).

In the first place, the mathematician and population geneticist R. A. Fisher (1918) devised a novel statistical method to account for those characteristics in which many Mendelian factors contribute to the phenotypic outcome. Despite his use of biometrical results, he rejected some fundamental assumptions made by Pearson (1904). He ignored Pearson's rejection of Mendelian dominance, and assumed that, in the combination of a large number of factors, the heterozygote could assume any value between the recessive and the dominant (thus rejecting complete dominance). As Mayo (2008: 252) highlights, in his 1918 paper, the HWP "was the basis for Fisher's derivation of correlations between related individuals under Mendelian inheritance". Even though he did not reference Hardy, also in his *On the dominance ratio* (1922) Fisher assumed that the distribution of the frequency ratio for different Mendelian factors could be determined from the stability of the distribution in the absence of selection, random survival effects, etc.[152] The assumption of equilibrium proportions was required also in the development of Fisher's concept of 'balanced polymorphism' (Mayo 2008), and allowed him to demonstrate – in his book *The Genetical Theory of Natural Selection* (1930) – that a marginal selective advantage for one allele was sufficient for this allele to spread through the entire population over time (Rheinberger & Müller-Wille 2017).

In the second place, as early as during his studies on the inheritance of coat colour in mammals, Wright (1917, 1918) assumed the HWP as equilibrium principle – even though he was not aware, at that time, of the work of Hardy and Weinberg – because this theoretical tool allowed him to formulate the appropriate genetical hypotheses about the genotypic distributions in populations. As noticed by Provine, such an assumption should have seemed rather self-evident to Wright, since he referred to it as to "the well known formula for a Mendelian population in equilibrium" (Wright 1917: 522, as cited in Provine 1971: 156). Nonetheless, later in his work his attention to complex gene-gene interactions and his sensitivity to the interests of animal breeders made him aware of certain unrealistic assumptions behind the mathematical modelling, especially in Fisher's simple models, and led him to regard them as purely instrumental (Rheinberger & Müller-Wille 2017).

Finally, J. B. S. Haldane (1924) used the HWP as the basis for all his important early work on selection in populations. He derived the recurrence relationship for the approach to equilibrium of gene and genotypic frequencies of an X-linked diallelic gene. Furthermore, Dobzhansky (1998) points out that Hardy's derivation played a fundamental function not only in the Anglo-American context, but also

[152] See Morrison (2007: 320-329) for a synthetic but comprehensive account of the trajectory that Fisher followed from the 1918 paper to the 1922 paper.

in the population genetics of the former Soviet Union, with specific reference to the work of Chetverikov.

In the light of these considerations, it seems fair to say that the HWP was held fixed across different empirical inquiries and theoretical developments during the early days of population genetics, in its epistemic function of equilibrium state for genetic populations. This fixity enabled the framing of the relevant questions to be asked and the testing of other theoretical components under development, especially in the absence of an established common framework of inquiry shared by the scientists. All these elements indicate a high degree of quasi-axiomaticity of the HWP already in the early days of population genetics, which further consolidated with its establishment as a discipline.

5.2.3.2 Generative potential

At a first glance, the HWP may appear as a sheer consequence of Mendel's laws of inheritance. Punnett's question to Hardy already encapsulates the connection of the HWP to Mendel's laws: why is it the case that the dominants do not eliminate the recessives in the long run? According to Edwards (2008: 1144):

> The answer is immediate from Mendel's first law. Segregation is independent of the segregants. Dominance has nothing to do with it […] it is an immediate consequence of Mendel's law of segregation that the expected frequencies of the genes among the offspring of two parents are equal to the frequencies of those genes in the parents themselves.

Yet, the HWP had a significant impact, such that the occurrence of its first explicit formulation is often made to coincide with the origins of population genetics (Crow 1988). The innovation brought about by the HWP consisted in generalising Mendel's experimental results and show their independence from any assumption of a specific initial distribution of gametes: "While Mendel conceived the independent binomial sampling of gametes from parents and hence could be regarded as the first to have considered a population–genetical example […] the generalization to arbitrary gene frequencies to give HWE was the true foundation of population genetics" (Mayo 2008: 254).[153] Rather than being just an empty law or a mathematical derivation that greatly simplified the mathematical treatment of the distribution of genetic variability, the history of the HWP shows that it performed a

153 As Edwards (2008) points out, the fact that Mendel's breeding experiments always started from cross-fertilising hybrids (*Aa*), so that the allele frequencies were indeed one half, plausibly had put Yule and Pearson "under the misapprehension that implicit in the Mendelian theory was the assumption that the gene frequency was one half, for then indeed a 3:1 ratio would appear and be maintained" (Edwards 2008: 1145), even though this assumption had no bearing on the actual frequency of alleles in a real population. This assumption pre-empted both Yule and Pearson from reaching an explicit formulation of the Hardy-Weinberg principle.

distinctive epistemic role. It demonstrated something about the preservation of genetic variability in a population, that is, the fact that there is no intrinsic tendency for genetic variation (in its simplest form, of the three different genotypes, one heterozygote and two homozygotes) to disappear in the long run (Morrison 2007).

Still, it might be argued that its epistemic role is that of a definition resulting from a mathematical demonstration and, as such, *a priori*, rather than attributing any conceptual 'generativity' or justificatory role to it. In other terms, one could argue that the HWP explicated certain relationships between assumptions of the Mendelian explanatory scheme, explication that was historically functional to enhance the development of population genetics, while the conceptual 'burden' remained on Mendel's laws.

Certainly, the HWP is not an empirical law or generalisation. Rather, it is better understood as functioning like a 'contrafactual ideal situation', where any deviation from the ratios it expresses points to hidden interactions in need of explanation (Rheinberger 2013: 488). In fact, its mathematical formulation does not enable any inferences to phenomena in and of itself, since the proportions it expresses can be subject to multiple interpretations, including a purely geometrical one (Millstein et al. 2009). On the contrary, its counterfactual character comes to work when the formalism is used to interpret a frequency distribution and to justify hypotheses on the causes of the deviations from expected frequencies. In other words, as a piece of counterfactual reasoning, the HWP posits an 'impossible' world or scenario against which real changes in genetic frequency distributions can be compared. Therefore, it is in virtue of its counterfactual character that the HWP justifies a mapping between a formal abstract representation and the concrete empirical phenomena. At the same time, this counterfactual situation represents a (re)conceptualisation of certain relationships between Mendelian assumptions under certain conditions as capturing the equilibrium state of (diploid and sexually mating) genetic populations *in general*. In this sense, the HWP came to represent the enabling condition to establish which real situations need an explanation, in that they deviate from that ideal situation conceptually determined as equilibrium state. As an upshot, it might be said that the HWP has a relatively high generative potential, but it does not approach a maximal degree, given that its 'generativity' within the domain of population genetics does not apply to all genetic populations (i.e., asexually mating populations and haploid organisms are outside of its domain of applicability).

5.2.3.3 Empirical shielding

I highlighted in the previous subsection that the HWP is not an empirical law or generalisation. As it is evident from the formulation of the HWP that I provided above, certain explicit assumptions enter the definition of the HWP, and determine its domain of applicability. The conditions under which a Hardy-Weinberg distribution holds cannot be found in nature, since there are no natural populations that are unaffected by evolutionary processes. In addition, Hardy-Weinberg proportions might hold even under the influence of evolutionary factors, e.g., when mutation and selection compensate one another. Consequently, the binomial distribution expressed by the HWP under its explicit conditions could not be directly tested anyway. This is, at least in part, a consequence of the highly idealised nature of the HWP. For instance, the requirement that the population be infinite for Hardy-Weinberg proportions to hold is motivated by the necessity to rule out statistical error, which approximates zero the more a population tends to infinity.[154]

Clearly, simplifying assumptions such as that of an infinite population are unrealistic, in that they posit 'impossible' worlds in which certain conditions might hold. Still, the gain in simplicity brought by the HWP as a theoretical tool is greatly advantageous from a pragmatic point of view: "[…] such models are useful because they strip a process to its essence and allow scientists to test particular attributes of a system in isolation" (Russell 2010: 604). Formulating hypotheses about the evolutionary factors that determine the different patches of genetic variation within and between populations seems to require the assumption of a state under which no such forces are active. However, additional assumptions beyond the absence of forces are also required, on the one hand, to model the equilibrium state according to the specific domain under investigation (for instance, assuming infinitely large populations reduces the risk of sampling error), on the other, to test for departures from equilibrium proportions (for instance, it is assumed that small departures from the assumptions lead to small departures from Hardy-Weinberg equilibrium, which is not trivial). In sum, these considerations indicate that the HWP has a high degree of empirical shielding.

5.2.4 Is the HWP constitutive?

In the light of the analysis I just provided, the HWP's epistemic function within population genetics could be indicatively represented in terms of its quasi-axiomaticity (QA), generative potential (GP), and empirical shielding (ES), as depicted in fig. 5.1. This graphic representation certainly leaves out many theoretical and conceptual components relevant to population genetics, such as Mendel's laws and the statistical tools deployed by population geneticists. Its aim is merely to display the relationship between the HWP, changes in genetic frequencies distributions, and their representation as deviations from expected equilibrium frequencies. Of course, a change in genetic frequency distribution is *not*

154 In actual practice, this requirement holds looser because statistical error is considered irrelevant for very large populations (usually $n > 500$).

constituted in and of itself, as it is obtained, for instance, through gene counting. But its conceptualisation as a deviation from equilibrium, and the outcomes of the explanatory practices implemented to account for such a deviation, are determined by what we take to be the equilibrium condition. Let us take a change in the distribution of one allele in a genetic population from time t_1 to time t_2 observed through gene counting, C, as our phenomenon. We attribute to it no quasi-axiomaticity, since it is an observation with no fixed epistemic function, no generative potential since, as a phenomenon, it does not have any conceptual generativity, and no empirical shielding, since it is obtained through empirical observation.[155] Given the HWP, i.e., the principle that provides the conditions for the equilibrium state of genetic populations – which, however, cannot be attained in real populations – C will be represented as a certain deviation from the expected frequencies provided by the HWP. This makes it possible to deploy it as a tool to investigate the action of evolutionary forces on a genetic population, since the extent of the departure of C from the ratios given by HWP provides an indication of the extent to which evolutionary factors are in action. Therefore, the quasi-axiomaticity of the representation of C as a deviation from Hardy-Weinberg ratios, call it D, increases compared to that of C, since D acquires a more distinct epistemic function of enabling the investigation of evolutionary causes in action. However, the generative potential of D does not increase compared to C, because it does not sustain itself any more conceptual dependencies. Its empirical shielding increases only slightly, since its representation as a deviation from expected frequencies implies some abstraction, provided by the HWP.

[155] As I explained in chapter 2, I attribute a 0 value to all three features to frame something as a phenomenon in the cube. The choice of what counts as a phenomenon is conventional in character and requires abstraction and idealisation, given the great complexity of scientific inquiry. The choice of what counts as a phenomenon is made to highlight the positions of the epistemic components of interest relative to one another and to a context of inquiry, not to point to absolute relationships between them. In this example, I only take into consideration the relationship between a concrete change in the distribution of an allele in a population, the HWP, and the representation of that frequency change as a deviation from expected frequencies in the context of contemporary population genetics. Among many things, I abstract away, for instance, from the fact that the method of gene counting as an observational method is not theory-free and, therefore, does imply a certain amount of empirical shielding.

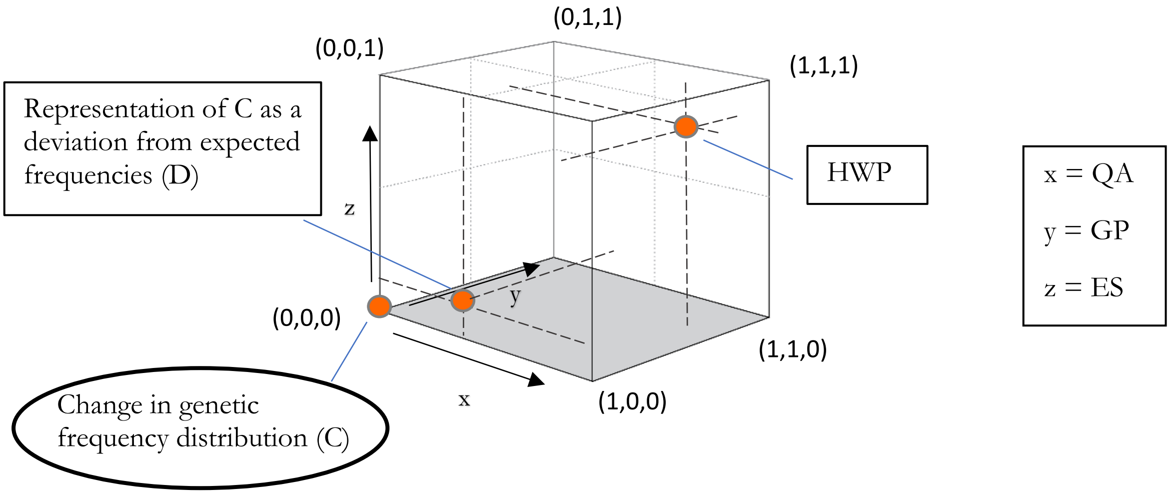


fig. 5.1: Graphic representation of the constitutivity of the HWP with respect to its epistemic function in the representation of changes in genetic frequency distributions as deviations from expected equilibrium frequencies

The HWP enables the representation of changes in frequency distributions as deviations from expected equilibrium ratios in order to investigate the evolutionary causes in action. In this sense, it provides part of the coordination for the representation of genetic frequency distributions as departures from equilibrium ratios.[156] Its quasi-axiomaticity is high, since it is held fixed to fulfil the epistemic function under analysis, i.e., that of providing the equilibrium state against which real genetic distributions are compared. As I discussed above, the HWP has generative potential in virtue of its counterfactual character. Yet, although it is crucial to conceptualise a large class of real frequency distributions as deviations from expected frequencies, the HWP does not provide the equilibrium conditions for all genetic populations. Therefore, its generative potential is high but not maximal. Finally, the HWP has high empirical shielding, since it expresses some abstract mathematical ratios and recurs to strong idealisations, but it does not approximate a maximal degree, since it also expresses certain physical requirements that must hold for those ratios to occur. Its high empirical shielding cannot simply be explained away by its interpretation as *a priori*, in Sober's sense. Even though the HWP may be definitionally *a priori*, this does not make it empirically shielded at the same level of, say, a mathematical language or a generative grammar, e.g., as the infinitesimal calculus in Newtonian mechanics. Indeed, the calculus *qua* mathematical language does not have an empirical meaning and,

[156] It is interesting to compare the coordination to which the HWP contributes with the one involving Newton's laws of motion that I discussed in chapter 2. The HWP is involved in the coordination between genetic frequency distributions and their representation as deviations from expected ratios in such a way that it is more constitutive than both. In the case of Newtonian mechanics, the coordinating component, i.e., the laws of motion, are less constitutive than one of the elements it coordinates, i.e., the infinitesimal calculus. This may look like a striking consequence of my account, but it results from my extension of the notion of coordination beyond Friedman's rigid three-level view of scientific theories.

thus, cannot be subject to confirmation or disconfirmation in and of itself. In the same way, the mathematical proportions expressed by the HWP do not have an empirical content of their own, but they acquire some, even though it is highly shielded by strong idealisations, in virtue of the counterfactual conditions expressed by the HWP.

Although historically the HWP has been crucial with respect to its function of providing the equilibrium conditions for genetic populations, both biologists and philosophers have explored plausible alternatives to it (Bosco et al. 2012, McShea & Brandon 2010). Certainly, the presence of empirical irregularities constitutes a reason to investigate these alternatives.[157] Still, their emergence would not lead to an outright disconfirmation of the HWP which, even in case of replacement, would still work as an appropriate modelling tool for a wide variety of purposes. The possibility of replacing or supplementing the HWP, in fact, does not depend on any direct empirical test of the HWP itself. It rather depends on the presence of competitive counterfactual alternatives that can fulfil the same epistemic function, but conceptually determine the domain of its applicability (or a larger one) in a different way. For instance, Bosco et al. (2012) suggest that the HWP as an equilibrium principle has important limitations that constrain its usefulness. By means of a numerical simulation, they show that, by assuming Hardy-Weinberg proportions, neither random nor non-random mating can be inferred for any finite population. Given the strong limitations of statistical tests to identify and characterise the equilibrium states of genetic systems, they argue that the very notion of equilibrium state as identified by the HWP would only be attainable in the infinite time horizon, and not by looking at isolated observations on a genetic population, which can, at most, capture 'stable' states. Consequently, they call for a reconceptualisation of the notion of equilibrium in population genetics and argue that the equilibrium state of genetic systems should be defined in terms of dynamical systems, by integrating a time variable to account for the evolution of the allele frequencies of the system over time. Another proposal is that of McShea & Brandon's (2010) zero force evolutionary law (ZFEL), which identifies the equilibrium state of evolutionary systems with their increasing state of diversity and complexity, rather than with a no-change state as the one captured by the HWP.

These alternatives to the HWP fulfil its same epistemic function, i.e., that of providing the equilibrium state to analyse evolutionary effects on Darwinian populations. However, by defining the equilibrium

157 Following up on the work of Li (1988), Stark (2006a, 2006b, 2007) showed that the assumption of random mating is sufficient but not necessary for Hardy-Weinberg proportions to hold in a population, but they can be reached also with non-random mating. Thus, even though experience has shown the fruitfulness of the HWP as a basis for hypothesis formulation and testing, this does not constitute a good reason for retaining the assumption of random mating (Stark & Seneta 2013).

state in different ways, they 'constitute' alternative representations of its domain of applicability, since which parameters – and relationships between them – within the domain of interest are considered as salient in the alternative representations is at least partially dependent on how the equilibrium state is characterised. Consequently, this aspect bears on what type of causes scientists look for when they develop their explanations, what further parameters they introduce for their modelling practices, etc. For instance, an equilibrium state characterised in terms of dynamical systems, as envisioned by Bosco et al. (2012), configures the system of interest (a certain genotypic distribution) as an entity developing over time, subject to many perturbations, and whose equilibrium condition can only be determined (asymptotically) in the long run. On the other hand, McShea and Brandon's attempt to identify an equilibrium state that could apply to all evolutionary systems, in contrast with the HWP's limited applicability, also has substantial consequences on the way the system of interest is modelled, and explanations are developed. Indeed, "[a] biologist armed with the ZFEL looks to stasis, rather than change, as the sure indicator that an evolutionary force is at work" (Gouvea 2015: 369). This is because they conceive of change itself as the state of equilibrium and, thus, the lack of deviations from Hardy-Weinberg proportions over time in a population would plausibly indicate, in their framework, the action of some evolutionary force to preserve an 'unnatural' condition of stasis.

In my view, these considerations highlight some aspects that cannot be fully captured by understanding the HWP simply as *a priori* in Sober's sense, even though it can be compatible with it. Certainly, the HWP *defines* the equilibrium conditions of genetic populations. However, it does so because it is an instance of counterfactual reasoning, which has a constitutive character in that it provides a justification for the coordination between some abstract representations and concrete phenomena. The HWP provides expected frequencies, against which real gene distributions are tested, thus it does not make up the real gene distributions themselves. Yet, as I highlighted above, their characterisation as deviations from equilibrium, and the outcomes of the explanatory practices implemented to account for these deviations, are determined by what we take to be the equilibrium condition. However, once we abstract away from the kind of 'generativity' that characterises the HWP, it seems clear that it is not maximal within population genetics, given the restricted domain to which it applies (diploid organisms, sexually mating).

5.3 The epistemic role of approximation and stability in relation to the HWP

5.3.1 Idealisation and the principle of approximation

As I emphasised above, idealisations are central to the formulation of the HWP. Yet, from the highly idealised HWP scientists construct less idealised and more empirical models of genetic populations. In this section, I focus on the source of epistemic justification for de-idealising from the HWP to make inferences regarding real populations.

Idealisations are systematic attributions, omissions, or distortions of certain features of a target domain for the purposes of simplification in modelling or representational activities.[158] In section 4.3.4, I mentioned that, according to van Fraassen (2008: 52), the use of idealisations in science requires assuming the 'principle of approximation', which he defines as follows: "If certain conditions follow from the ideal case, then approximately those conditions will follow from an approximation to the ideal case". In other words, the possibility of constructing more empirically accurate representations $X_1,\ldots X_n$ of a system from a highly idealised representation of it, Y, presupposes that the dependent variables in the representations $X_1,\ldots X_n$ will approximate the ones in the idealised representation, if the empirical parameter(s) in $X_1,\ldots X_n$ approximate the idealised feature(s) in Y. Someone could rebut that, since idealisations are widely used in science and 'approximation' is a vague term, this principle is trivial and, perhaps, that it says nothing about how to de-idealise from the HWP in modelling practices of population genetics. In fact, van Fraassen himself highlights that local background considerations are essential to understand when and to what extent the principle of approximation holds. Practices of de-idealisation are subject to local considerations, dependent on the theoretical context and on the run-of-the-mill pragmatics of model-tweaking. The principle of approximation, however, is not trivial. The fact that – in its most general form, as the one provided by van Fraassen – it is taken for granted does not mean that it is empty, but rather demonstrates its domain-independence, given that idealising practices are pervasive in the sciences. This is even more important once we consider that the use of idealisations is itself often taken for granted, and the specific human contribution to idealising as an epistemic activity is not always recognised as such, especially with respect to those domains in which talk of 'laws' and 'exactness' is dominant (Teller 2001, 2004).

In the case of the HWP, we can see the principle of approximation in action in the assumption that small departures from the ideal conditions expressed by the HWP lead to small departures from Hardy-Weinberg proportions. This assumption, as I have argued, is not trivial even if it is taken for

[158] Cf. above, chapter 3, section 3.3.1, especially footnote 87.

granted in the modelling practices of population geneticists. A particularly striking case is the requirement of infinite populations for Hardy-Weinberg proportions to hold. This requirement is supposed to rule out the effects of drift which, in population genetics modelling, is identified with statistical error. If a population is infinite, the error rate coming from sampling is zero, but for practical purposes this requirement holds looser. If a population is large enough (for most purposes, > 500 individuals) statistical error is considered irrelevant. Therefore, the principle of approximation underlies the following inference: if a population approximates infinite size, then the effects of drift are approximately zero.

What about the requirement of infinite populations itself? Even if its use is epistemically justified, we might still wonder what it refers to. According to Abrams (2006: 265), the requirement of infinite populations should not be taken as referring but only as exhibiting boundary limits, given that "a claim about an infinite population is a shorthand for a claim about a limit as the population size is increased without bound". This interpretation is very much consistent with the interpretation of the HWP as a counterfactual ideal scenario, since it emphasises that the HWP does not, and cannot, refer to any real population, but it can still represent ideal populations by means of its counterfactual character. On top of this, the principle of approximation underlies the use of the HWP in the construction of more realistic models, thus constituting a presupposition for its empirical application. In short, the principle of approximation is involved in the coordination between two abstract representations, the highly idealised HWP and the less idealised models developed by de-idealising from it. Thus, the principle of approximation is constitutive because it is required to justify inferences from the HWP to more empirical models, and not directly to phenomena. Clearly, the principle of approximation is not a domain-specific constitutive component, since it justifies the use of idealisations and underlies the practices of de-idealisation across most scientific frameworks, as I have shown in chapter 4 with respect to its role in calibration and measurement.

5.3.2 Equilibrium states, the principle of stability, and counterfactual reasoning

In section 5.2.4, I mentioned how the possible alternatives to the HWP that have recently emerged might 'threaten' its state of (unique) equilibrium state for genetic populations. This, in turn, would have repercussions on the available conceptualisations of what counts as a deviation from the equilibrium state and the strategies to account for these deviations. Given these considerations, an important question could be raised: how can we establish when the use of the HWP *qua* equilibrium state is epistemically justified?

Again, I would like to stress how local background considerations and the purposes of the epistemic subject(s) can be crucial in determining which one among the available alternatives is the most suitable for a specific experimental or modelling goal. Yet, the emergence of equilibrium principles alternative to the HWP indicates that there might be a more abstract and epistemologically significant notion of equilibrium that encompasses all of them. In this sense, some general considerations on the role of 'stability' as a notion can be illuminating. The HWP as an equilibrium state expresses the idea of a stable system, whereby it identifies a single distinct condition of balance of a genetic population, even though this configuration is only possible, strictly speaking, counterfactually. Consequently, we might wonder whether stability is an epistemic feature of our representations, an ontological feature of allelic distributions, or both.

I have previously illustrated how Bosco et al. (2012) argue that the equilibrium condition, as the state in which genetic populations are not changing, can only be identified by considering their development in the infinite time horizon. According to them, this condition cannot be inferred only by looking at isolated observations of a genetic population and comparing them to the idealised prediction offered by the HWP. These isolated observations can, at most, capture stable states of populations, in which there is an (approximately) invariant distribution of alleles. Therefore, they call for a reconceptualisation of the equilibrium state of genetic populations in terms of dynamical systems, by integrating a time variable to account for the evolution of the allele frequencies. This reconceptualisation brings along also the question of what we *mean* by a stable state. Therefore, the identification of what counts as a stable state in a domain seems to be a general prerequisite for representing a dynamical system.

Within dynamical systems theory the notion of stability encompasses several distinct mathematical relationships between system states and phase portraits – i.e., sequences of states under a certain dynamical rule – and systems themselves can be in equilibrium states characterised by 'stabilities' (or 'instabilities') of different sorts (Holmes & Shea-Brown 2006). A general characterisation of 'stability' in dynamical systems theory could be given by saying, in rough terms, that a system is stable if its mathematical description in terms of its phase portrait will remain in the topological neighbourhood if the system is subject to perturbations, i.e., the initial conditions are altered.

One of the more specific meanings of 'stability' in dynamical systems theory is that of 'structural stability', which is supposed to capture whatever is preserved of the structure of a system when a system is under perturbation, even if microscopic details change (Pugh & Peixoto 2008). According to Batterman's (2001: 58) definition,

> The idea of a structurally stable solution is, roughly speaking, the idea that because the coefficients of the differential equations we take to govern the behavior of the phenomenon of interest are never determined precisely, the *global structure* of the solutions to those equations ought to be relatively insensitive to small changes in the forms or values of those coefficients.[159]

Nikolov et al. (2007) offer a characterisation of 'structural stability' that could be applied to biological systems, as a property of a system which preserves its structure unaltered, i.e., it does not change the number and the type of (i) equilibrium states or (ii) periodical movements. This definition differs from their own characterisation of 'stability' *tout court* as property of a system that, if under perturbations, remains in one and the same state. This distinction is supposed to mirror the one between stability and structural stability in the mathematical sense of dynamical systems theory that I explained above.

The equilibrium condition of a genetic population described by the HWP is that of the only zero-force state of a system, that is, the state at which the system is supposedly unchanging, since there are no perturbations due to the action of evolutionary factors. This is very different from what the notion of 'structural stability' seems to express, according to both definitions illustrated in the previous paragraph. As Bosco et al.'s research emphasises, considering the HWP as the only stable equilibrium state for genetic populations requires the assumption of a host of ceteris paribus conditions and, most importantly, requires leaving out time-development as a parameter internal to the system under consideration. According to them, a genetic population can be understood as being in a structurally stable equilibrium only if its macro-properties are robust across time, even under perturbations. However, according to the HWP, a population can be in equilibrium condition only at a time, t_0, and this state can be preserved at time t_1 only if the initial conditions are unaltered, i.e., if no evolutionary factor is in action. In other words, the equilibrium state expressed by the HWP does not seem to be a stable one, since even small perturbations alter the ratios it predicts and, indeed, the conditions that it prescribes are attained only in a highly idealised scenario. Therefore, structural stability in the ontological sense, as a feature of real systems does not seem to be involved by the HWP as expressing the conditions for the equilibrium state of genetic populations.

Yet, stability in an epistemological sense could be doing some work in justifying the use of the HWP as equilibrium state. Fletcher (2017) has suggested an epistemic interpretation of a notion of stability that has a substantial role in representational modelling. More precisely, stability can be conceived as a necessary epistemic assumption required by our practices of representational modelling to justify our inferences from models to phenomena. According to Fletcher (2017: 2): "an inference from a

[159] Italics in the original text.

property of a model to the property of phenomena (or the world) it represents is justified only if all sufficiently similar models have that property".[160] This principle of stability regulates the attribution of properties of a model to the phenomenon or system that the model is supposed to represent, by prescribing that such an attribution is legitimate only when similar models have the same property, where the similarity relationship can be characterised in topological terms. Fletcher's principle of stability is different from the principle of approximation described above, because the latter concerns the relationship between two models or representations, where one can be more or less idealised than the other, while Fletcher's principle of stability concerns the relationship between a class of models and a phenomenon. At the same time, the principle of stability is also different from what Woodward (2006) calls 'derivational robustness', i.e., that feature of a model by which, if we change a value of one of its parameters, we can still derive the same or a very similar result from it. The most notable difference between Fletcher's principle of stability and Woodward's robustness is that the former is an assumption concerning the justification of inferences from models to phenomena, while the latter is a virtue of a model that comes in degrees.

According to Fletcher, a clear example of his principle of stability comes from the epistemological understanding of 'structural stability', as characterised in dynamical systems theory. A dynamical system is structurally stable just when its structure is mostly preserved under perturbations, that is, when all systems with sufficiently similar dynamical rules provide qualitatively similar phase portraits. In other words, if there are any fixed points in the phase portrait of a structurally stable dynamical system – states at which the system is unchanging under the dynamical rule – they and their qualitative features will be the same for all sufficiently similar dynamical systems. As Fletcher highlights, there seems to be a consensus in dynamical systems theory that this is an *epistemic* constraint, rather than an ontological one. However, Fletcher generalises from this example and extends the epistemic role of the principle of stability well beyond the domain of dynamical systems theory. According to him, the principle of stability regulates the possible inferences from classes of models to concrete phenomena *in general*.

I will not argue here for or against Fletcher's own general characterisation of the stability principle. More modestly, I side with Fletcher's understanding of 'structural stability' in dynamical systems theory as a good example of his epistemological principle of stability, which identifies a precondition for modelling dynamical systems and, therefore, genetic populations in Bosco et al.'s sense. If we were

[160] Fletcher attributes the first formulation of the principle of stability as he characterises it to Duhem (1914/1954).

to use Bosco et al.'s (2012) characterisation of equilibrium state to model frequency distributions of genetic populations, we would see in action an idea of stability that has the same epistemic function as described by Fletcher in the context of dynamical systems theory. In this case, the ratios expressed by Hardy-Weinberg proportions under the well-known conditions would not represent *the* equilibrium state of a genetic population. Instead, they would represent only one of the stable equilibrium states in which a genetic population could be, where such a stable state is one in which statistical invariance of allele distribution holds. This means that the long-term behaviour of the genetic population under consideration (measured in terms of the asymptotic time invariant values of the allele frequencies) can present strong variations, and what counts as *the* equilibrium condition for that population can only be assessed asymptotically in the infinite time limit. The idea of 'structural stability' as an *epistemic* constraint is here at work since, even though there could be a great deal of micro-scale variations (i.e., in the allelic distributions at single times $t_1 \ldots t_n$) in the development of two systems X and Y, if the stable equilibrium states are preserved in time-evolution, this will justify X and Y having a similar description.

On the contrary, as I highlighted in the beginning of this chapter, the sense of stability that is often associated to the HWP as equilibrium state in classic population genetics appears to be linked to the projection of a stable state as 'inertial'. However, the stability expressed by the HWP in classical genetics does not denote an equilibrium in the mechanical sense; rather, it refers to the statistical invariance of allele frequencies. For this reason, Hardy, in his 1908 paper, carefully chose the term 'stability' over 'equilibrium' (Mayo 2008). Despite the attractiveness of a more metaphysical understanding of equilibrium in the mechanical sense, by which HWP would be the resultant of the null pressure of evolutionary forces,[161] the sense of stability at work in the use of the HWP as equilibrium state concerns its representation of *the* counterfactual state in which statistical invariance of allele distribution holds. Any additional metaphysical weight on the supposed 'mechanical' character of this equilibrium condition comes from further ontological or causal assumptions, usually based on analogies with classical mechanics. Yet, the epistemic role of stability does not seem to concern the preservation of the equilibrium state under perturbations, but rather the picking out of a single identifiable condition in which the system is not changing, i.e., where the only equilibrium condition is univocally determined by certain conditions. However, even though the equilibrium cannot be preserved unless those conditions hold, 'quasi-equilibrium' states can hold under similar conditions. Therefore, the idea of stability here seems to constrain the set of possible states in which invariance

[161] Cf. above, section 5.1.3.

of allelic frequencies approximately holds and, at the same time, it incorporates the assumption that there is only one single equilibrium state for genetic populations, identified by the Hardy-Weinberg conditions.

Although it is hard to generalise from only two examples concerning the characterisation of the equilibrium condition in population genetics, I believe that some interesting conclusions can be drawn concerning the epistemic role of a general 'principle' of stability in coordination. In my view, these examples show that a rather general notion of stability is at work in justifying certain relationships between representation and phenomena in close connection with counterfactual reasoning. Both examples, i.e., the HWP as equilibrium condition and the alternative characterisation of equilibrium in a dynamical sense, involve two different specific senses of stability. Yet, in both cases it is presupposed that a model or class of models is stable *in general*, in the sense of providing one or more relatively fixed points, in the form of equilibrium conditions, such that population geneticists can make inferences relative to an empirical domain which extends well beyond the specific state or set of states described by that model itself. In the case of the HWP, stability has the form of an epistemic constraint to the range of applicability of the model representing the equilibrium state, such that it reflects the presence of a single equilibrium condition. This latter aspect has, in some cases, also led to a misleading understanding of the HWP in a mechanical sense, as the resultant of null evolutionary forces, given that under this interpretation there can be only one equilibrium state, understood as physical absence of change (rather than statistical invariance). However, this step is not justified by the idea of stability itself, but should be justified through a form of coordination between the formalism expressed by Hardy-Weinberg proportions and the specific concepts of drift and fitness.[162] In the case of the dynamical systems interpretation of equilibrium for genetic populations, the idea of stability functions, again, as an epistemic constraint justifying the applicability of similar descriptions to genetic populations that preserve their overall macro-structure and, therefore, their equilibrium states, along time-development. In addition, in both cases the stability assumed with respect to the model(s) identifying the equilibrium conditions is essential to make inferences concerning real frequency distributions as deviations from those equilibrium conditions.

The difference with the principle of approximation seems subtle but it is crucial. Approximation concerns the relationship between two representations, whereby if from an idealised representation

[162] Most plausibly, this coordination is justified also by the measurement procedures used to identify drift and fitness as quantities, on similar lines to what I discussed in chapter 4. Due to pragmatic constraints, I cannot delve into this complex issue here.

certain conditions follow, then from an approximation to the idealised representation approximately those conditions will follow. On the contrary, stability concerns the relationship between a representation or a class of representations and phenomena, such that if a representation is stable, i.e., it does refer univocally to a real or ideal state or class of states, similar representations will appropriately refer to similar states. It follows that stability is a crucial assumption that enables the investigation of empirical domains by varying certain features of a representation, e.g., the values of a parameter in a mathematical model. This is typical of several epistemic practices involving counterfactual reasoning, such as thought experimentation and causal modelling, where the epistemic subject(s) 'manipulate' features of a representation to infer the possible consequences of those variations on concrete states. In addition, this is presupposed by real experimentation, if we consider counterfactual reasoning as a requirement for actual experiments, too, on Buzzoni's lines.[163] For instance, in order to vary experimental conditions and control for possible hidden variables, which are identified by reasoning counterfactually on a possible or real experimental situation, we must assume the stability of the minimal form of representation of a phenomenon, i.e., its (re)production in an experimental setting. In this sense, stability is required to reason counterfactually in a wide range of epistemic activities that are involved in various forms of coordination between abstract representations and concrete phenomena.

5.3.3 Approximation and stability as domain-independent assumptions for reasoning abilities

In fig. 5.2, I plotted the principle of approximation and that of stability in graph shown in fig. 5.1. Even in this case, the graph should not be viewed as a faithful representation of the structure of population genetics. What is being represented here are the relationships between a domain-specific component, the HWP, two domain-independent components, approximation and stability, and what had already been framed in fig 5.1 as the phenomenon (C) and its abstract representation (D).

[163] Cf. above, section 5.1.2.

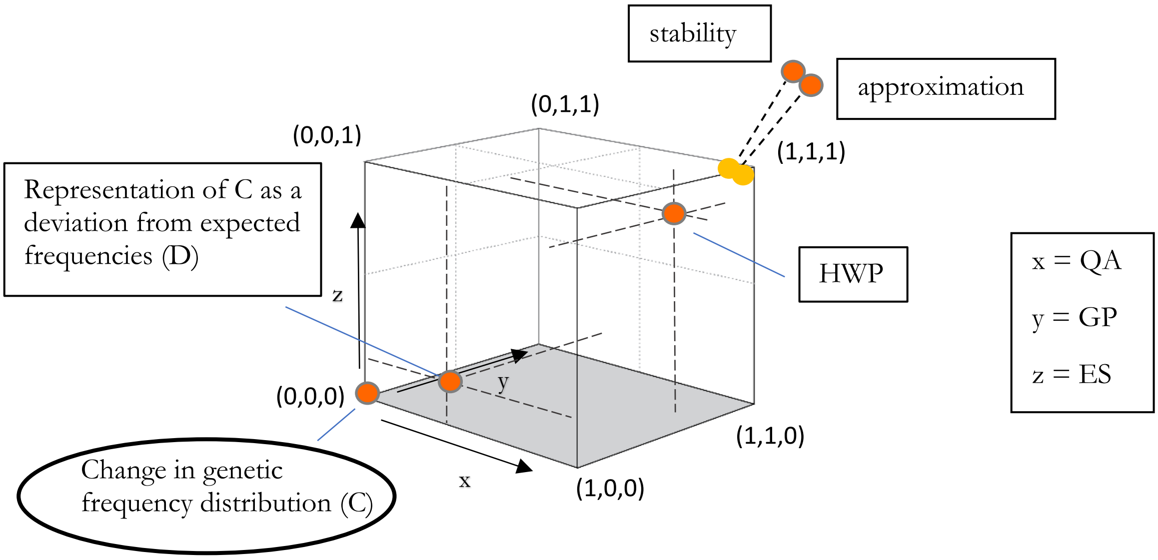


fig. 5.2: Graphic representation of the position of approximation and stability with respect to the HWP

The lighter circles represent the position of approximation and stability with respect to the HWP, a specific change in genetic frequency distribution, C, and its representation as a deviation from equilibrium, D. In other words, the positions of the lighter circles represent the extent to which approximation and stability have a constitutive character in the coordination between C and D *within* the framework of population genetics. Yet, they are constitutive in a domain-independent sense because they are preconditions which underlie several epistemic activities (idealising, modelling, thought experimenting, etc.) that scientists perform across scientific frameworks. The darker circles connected to the lighter ones by a dashed line represent the fact that approximation and stability are not specific to the framework of population genetics.

The principle of approximation and that of stability are not involved only in the coordination between changes in genetic frequency distributions and their representation as deviations from the equilibrium conditions via the HWP. For example, as I mentioned when discussing calibration as a modelling activity in Ohm's scientific practice in chapter 4, the principle of approximation was presupposed by that activity too, in that it provided part of the justification of the inferences from instrument readings to measurement outcomes. In the context of the HWP case, I showed that the principle of approximation is required to make sense of the idealisation of infinite populations expressed by the HWP so that it can be deployed to construct less idealised models representing changes in genetic frequency distributions.

In both examples, the principle of approximation underpins the process of de-idealisation and, therefore, can be characterised as an assumption that must be presupposed to idealise and de-idealise, which are reasoning abilities involved in forms of coordination across several, if not all, scientific inquiries. The principle of stability can be understood along similar lines, with respect to the ability to reason counterfactually that is required by many epistemic practices. As I have shown, a specific understanding of stability is required by both the population geneticists' practice of representational modelling that takes the HWP as a zero-force model, and by the alternative modelling practice suggested by those researchers who interpret the HWP as only one of the possible stable states of a genetic population. Yet, this example could be generalised, as Fletcher does, to suggest that all practices of representational modelling require assuming a principle of stability, according to which, if we want to attribute a property of a model to a concrete phenomenon, all sufficiently similar models must have that property. My suggestion, on similar lines, has been to understand the principle of stability as concerning the relationship between a representation or a class of representations and phenomena, in the sense that if a representation is stable, i.e., it does refer univocally to a real or ideal state or class of states, similar representations will appropriately refer to similar states. In addition, I have related this assumption to our ability to reason counterfactually, which is required by several epistemic practices. As I have spelled out in section 5.1.2, counterfactual reasoning is a widespread tool in scientific inquiry, required for the positing itself of non-realised scenarios for some epistemic purpose, including thought experimentation, prediction, and hypothesis-making, but also for real experimental design. Counterfactual reasoning appears to be necessary in all epistemic activities with a goal that involves the understanding of causal connections by ideal entities under ideal conditions. It is not serendipitous that the most influential philosophical account of causation in the sciences is a counterfactual one (cf. Woodward 2003). Even in real experimentation in the laboratory, which is about real entities under real conditions, counterfactual reasoning is crucial in that it enables the normative anticipation of the possible causal nexuses to be identified, thus justifying certain pragmatic interventions on the experimental setting to accommodate actual divergences from the ideal (imagined) situation. The assumption of stability, i.e., that if a representation refers univocally to a real or ideal state or class of states, similar representations will appropriately refer to similar states, is required to reason counterfactually. This assumption prescribes that a certain referential relationship between a representation (idealised or less idealised) and its referent (real or ideal) must be stable, if we want to investigate similar referents by means of similar representations, that is, if we want to investigate those referents by varying counterfactually certain features of the original representation. In the case of the HWP, the assumption of stability concerns the fact that, within classical population genetics, it univocally refers to the ideal state of equilibrium of a genetic population, in such a way that this enables to analyse all real (and ideal) variations from that state as departures from equilibrium and investigate the factors or causes behind it.

As for the epistemic status of these components, I suggest that they may be understood as domain-independent presuppositions that we must assume in order to use certain reasoning abilities for some epistemic purposes in one or more scientific frameworks. In the table below, I present only those discussed in detail in this dissertation (fig. 5.3).

Reasoning ability	**Domain-independent principle**
Idealising and de-idealising	Principle of approximation
Counterfactual reasoning	Principle of stability

fig. 5.3: table representing the pairings of domain-independent principles and reasoning abilities

These domain-independent principles are constitutive insofar as they are required to make use of those reasoning abilities that are involved in some form of coordination, that is, that provide part of the justification for the representation of concrete phenomena in terms of abstract conceptual resources. These epistemic components, considered in the abstract, are domain-independent, since they are tied to reasoning abilities, rather than to certain specific domains of inquiry. In the same way, they could be considered as theory-independent, because the reasoning abilities themselves are not reducible to theory-specific considerations. For instance, counterfactual reasoning can be deployed to anticipate certain real causal connections in experimental practices in which the theoretical background plays a minimal justificatory role, such as in exploratory experiments.[164] However, this does not mean that these general presuppositions are, or can be, assumed in complete isolation from local background theory. Quite the contrary: since the reasoning abilities they underlie are always performed for a certain epistemic goal and, therefore, against the background of a specific epistemic standpoint, domain-relative assumptions are always involved in justifying the performing of certain activities rather than others. In the case of idealising and the principle of approximation, for example, the extent to which a certain kind of idealisation can be deployed will most likely depend on at least some features of the conceptualisation of the specific target domain. It is in this sense that, with respect to the issue of coordination, these domain-independent components can be thought of as

164 Cf. above, chapter 1, section 1.3.3, and chapter 4, section 4.3.2.

independent, but they always share their role of coordinating abstract representational tools with concrete phenomena with theory-relative principles and, in some cases, with material conditions.

The table in fig. 5.3 can be fruitfully compared to Chang's (2008, 2009) list of pairs of epistemic activities and transcendental assumptions required for their performance, of which I provide a sample in fig. 5.4. As I discussed in chapter 1 (section 1.2.3.3), Chang's principles are assumptions that we must hold so that our epistemic activities can be intelligible for us. In other words, if we do not assume the validity of any one of these principles, the associated activity would be pragmatically impossible to perform, since it would not be intelligible for us.

Epistemic activity	**Principle**
Counting	Discreteness
Testing-by-overdetermination	Single value
(Rational) prediction	Uniform consequence
(Linear) ordering	Transitivity
…	…

Fig. 5.4: Part of Chang's (2008) list pairing epistemic activities with transcendental principles

Chang's table and mine may look, *prima facie*, very similar. Important similarities between his principles and mine include their transcendental character, the fact that they are devoid of empirical content (have maximal empirical shielding, in my terminology), and their independence from a specific domain of application, i.e., from a specific conceptualisation of a domain. However, there are several differences, too. First, the principles I discussed are not present in Chang's list. Secondly, he connects his principles directly with epistemic activities, whereas I pair them with the reasoning abilities that we use when pursuing some epistemic goal. Certainly, the reasoning abilities that I have discussed are themselves deployed to perform several epistemic activities, including some of those identified by Chang as the most 'basic' ones. Yet, that of 'reasoning abilities' is not an extra category that I arbitrarily place in-between the principles and the activities. On the contrary, given that these reasoning abilities must themselves be used to perform the activities, as I have shown for the cases of idealising and reasoning counterfactually with respect to several activities such as thought experimenting, modelling, etc., I consider them a more fundamental unit than the activities themselves. Finally, our principles differ with respect to the feature of domain-independence, in that the principles I discussed are theory-independent only in the abstract limiting case, while they are, in fact, always involved in forms of coordination together with domain-specific assumptions. In Chang's account, on the other hand, they seem to be entirely autonomous, with respect to domain-relative conceptualisations of an empirical domain, although he allows for combinations of more principles when complex epistemic activities are performed.

Epilogue

I. Summary

The core claim of this dissertation is that different epistemic components can be variously involved in coordination and, thus, be considered as constitutive to some degree within their scientific framework, or even across frameworks, if they are domain-independent. By 'coordination' I referred, in general, to both the process and the condition of justification of the referential relationship between empirical phenomena and the abstract tools that scientists use to represent them. In this work, I have also drawn a conceptual distinction between this broad characterisation of coordination and a more specific meaning of it. I labelled the latter 'measurement coordination' and showed how coordination of such smaller scope already involves several entangled epistemic dimensions and activities. By being 'constitutive', I meant that these components take on a peculiar justificatory role in the process through which a form of coordination between some phenomena and their abstract representation is achieved as well as in the state of stable justification that characterises a 'mature' form of coordination.

As we saw, the role of these epistemic components in coordination is similar across different contexts of scientific inquiry. However, a central goal of this work has been to account for the variability among these components with respect to certain pivotal quantitative and qualitative characteristics. On the one hand, I have suggested that different epistemic components can contribute to coordination to various degrees. More precisely, I have developed a quantitative account, based on three features of these components in order to express to what degree they have a constitutive character in coordination. The degrees of quasi-axiomaticity, generative potential, and empirical shielding of the components under analysis, compared to those of other components within the same scientific framework, enable to assess the extent to which they are constitutive of a certain form of coordination. Evidence for attributing these quantitative differences comes from analysing the history and practice of using certain specific abstract tools to represent concrete phenomena, as well as the justification that scientists rely on for doing so. However, as I specified in the course of this dissertation, meaningful comparisons by means of my quantitative tools can be made from within a framework and they can only provide us with ordering relations ('more constitutive than…' or 'less constitutive than…').

On the other hand, I have argued that there are at least three qualitatively distinct varieties of epistemic components that can take on a constitutive character in coordination. As a result of applying my quantitative features to three case studies, I have identified important differences in kind between what I described as domain-specific theoretical principles, material components, and domain-independent assumptions underlying reasoning abilities. The domain-specific theoretical principles are presuppositions that crucially contribute to the justification of the referential relationship between a theoretical term or definition and the phenomenon that it describes, or between an abstract law or principle and the empirical regularity that it expresses. These presuppositions have a structural role within the theory they belong to, since they provide certain conditions of applicability of those theoretical terms or laws to concrete empirical material. At the same time, they are themselves often the product of a switch from being inductive generalisations to deductive axioms, as I have shown in the case of the core Darwinian principles and in the case of Ohm's assumption of equal distribution of tension.

In some contexts of scientific inquiry, material components can take on a justificatory role within specific forms of coordination and, thus, be constitutive to some degree. In the case of Ohm's thermocouple, I have highlighted how his trust in the productive role of the instrument, despite limited theoretical understanding of its workings, was crucial to the identification of the quantity of exciting force as a measure of the intensity of the electric stream via a determined measurement procedure. In this case, justification does not rely purely on background theory, e.g., in the form of measurement laws, or on abstract principles, but is partly rooted in some material conditions of inquiry, more specifically, on some material features of experimental practice. Yet, the instances in which material features take on a constitutive character in coordination are plausibly limited to specific historical episodes and localised contexts of inquiry, since more theoretical and abstract forms of justification eventually tend to outweigh them when a discipline or research program matures. In this sense, these components are even more domain-specific than the theoretical principles I discussed above.

Finally, some very abstract presuppositions seem to underlie general reasoning abilities that are involved in establishing and justifying forms of coordination across empirical domains. The case of the principle of approximation demonstrates how these presuppositions can be transversally pervasive in science, and yet, so hard to isolate and identify according to their epistemic function. The principle of approximation underlies idealising and de-idealising practices in scientific contexts as different as Ohm's calibration in his experiments on conductivity and resistance, and the use of the Hardy-Weinberg principle to conceptualise frequency changes in genetic populations. The same clearly applies to the idea or principle of stability as underpinning the use of counterfactual reasoning for representational modelling. Although it is really tempting to interpret at least some of these presuppositions in a metaphysical sense, as very general assumptions about reality, a word of warning is due. According to Chang (2008, 2009), similar epistemic components are best interpreted as metaphysical, albeit in a transcendental sense: they must be assumed as features of reality, for us to perform certain epistemic activities.[165] What I want to emphasise is that even these very broad assumptions can have a different epistemic 'flavour' depending on the situated context of inquiry in which they take on a justificatory role in coordination. For example, when they are used in inquiries of under-explored empirical domains they might be best viewed as having more of a heuristic character. Yet, when they have a justificatory role for forms of coordination in well-established theories or experimental systems it often seems that their metaphysical weight is heavier. Again, a transcendental interpretation of these assumptions allows for more wiggle room to account for the fact that contextual factors, such as values and goals of the epistemic subjects as well as the historical context do indeed influence their metaphysical weight.

II. Main innovations of this dissertation

In this section, I would like to highlight the innovative aspects of my dissertation, especially in comparison to other perspectives on the (constitutive) *a priori* and coordination in scientific inquiry.

First, the account that I develop in chapter 2 reconceives the traditional opposition between interpretations of the *a priori* whereby it can have empirical content and contentless interpretations as well as, to some extent, the opposition between *a priori* and *a posteriori*. The extent to which an epistemic component is constitutive depends on its function of contributing to the justification of one or more forms of coordination. In my view, this peculiar function of an epistemic component need be assessed in the context of its relationship with – and in comparison with – other components of an epistemic framework. In this sense, my account is contextual, relational, and, most importantly, it allows for degrees. Depending on its degree of quasi-axiomaticity, generative potential, and empirical shielding in comparison to the other components of a framework, an epistemic component will be more constitutive or less constitutive of a certain form of coordination than them. An upshot of conceiving empirical shielding as a feature that comes in degrees is that I can reconcile the contraposition between interpretations of the *a priori* whereby it can have empirical content and contentless interpretations. Some constitutive components will have a very high or maximal degree of empirical shielding, thus being more akin to contentless views of the *a priori*, while others have less empirical shielding and will, therefore, be closer to the other sort of interpretation. At the same time, my account relativizes even further the distinction between (constitutive) *a priori* and *a posteriori*. Friedman's innovation, following Reichenbach, has been that of considering certain components as constitutively *a priori* in a scientific

[165] Cf. above, chapter 1, section 1.2.3.3.

framework because of their function in coordination and independently of their having empirical content. My account builds on this innovation by relativizing their constitutivity, on the one hand, to the three features that they exhibit to different degrees and, on the other, to factors depending on context and on the relationship of these epistemic components with others in a framework.

Secondly, in this work I have bridged the literature on the general problem of coordination between theory and phenomena with the flourishing literature focusing on coordination as a typical issue in epistemology of measurement. Whereas the former debate directly discusses the role of certain epistemic components as constitutive, the latter is more recent and resorts mostly to coherentist and empiricist solutions. In chapter 4, I have connected these two bodies of literature on coordination by distinguishing between two kinds or scopes of coordination, a more general one between theory and phenomena, and a narrower one between measurement procedures and quantities. This distinction between a general and a narrower scope of coordination is only an initial attempt to show that coordination is often highly entangled. In fact, what I suggest is that coordinating abstract representations with the empirical phenomena they are supposed to represent can involve several 'layers' at different levels of abstraction, rather than only one, as in Friedman's analysis of space-time theories (recall his distinction between mathematical, mechanical, and empirical levels).[166] Depending on the specific context of the scientific inquiry under analysis, including the epistemic tools available (mathematical, conceptual, material, etc.), the degree of maturity of the discipline, and the complexity of the phenomena investigated the number of layers comprising the coordination between the concrete phenomena and their abstract representation can vary. In addition, different epistemic dimensions intersect in coordination, including measurement, experimentation, modelling, theorising, instrumentation, etc. Yet, I have shown how the same quantitative tools can be applied to express the constitutive character of components involved in different scopes of coordination.

Finally, I have identified three qualitatively different kinds of constitutive components, that is, domain-relative theoretical principles, material conditions, and domain-independent assumptions underlying reasoning abilities. These distinct kinds can coexist in a unitary picture. Understanding that coordination in its general sense, between some abstract representation and the empirical phenomenon it captures, often involves several smaller-scale layers of coordination, where various epistemic dimensions intersect, leads by itself to the conclusion that different kinds of components can have a constitutive character by contributing to the coordination. This is because these distinct

166 Cf. above, chapter 1 section 1.2.2. See also Padovani (2011), who first suggested the idea of multiple layers of coordination.

kinds of components typically have a function in different but entangled epistemic dimensions: domain-relative principles in theorising, material conditions in instrumentation and experimentation, domain-independent assumptions in several reasoning patterns involved by modelling, theorising, experimentation, etc. Indeed, comparisons between qualitatively different components can be done to establish which of them has higher quasi-axiomaticity, generative potential, or empirical shielding within a framework. Yet, these comparisons cannot establish whether some kinds of components are 'paradigmatically' constitutive in general. Such a claim is not justified given the qualitative differences between these kinds, on top of the contextual and relational character of the comparisons. In sum, these qualitative differences seem to justify a pluralism with respect to the presence of distinct kinds of constitutive components.

In a nutshell, the *gradualist interpretation* of the constitutive character of some epistemic components in coordination and the *pluralist perspective* on different kinds of such components are the core innovations with respect to other views of the constitutive *a priori*. I have already discussed at length my contribution with respect to one of my main benchmarks, i.e., Friedman's view of the relativised *a priori*. Compared to Friedman, my view does not restrict the class of constitutive components to one single set of principles at a time, to a single kind of components (i.e., theoretical domain-relative), or to components belonging to a single scientific discipline (i.e., space-time physics). By detaching his interpretation of the constitutive *a priori* from his three-layered analysis of coordination in physical theories, I have fully developed the gradualism implicit in his view and suggested that coordination can involve many layers and, thus, several kinds of epistemic components.

Other philosophical accounts of coordination or of the (constitutive) *a priori* have influenced this dissertation. Yet, the core features of my perspective are innovative with respect to them. Following Stump, I have given up the label of *a priori* on top of the 'constitutive' one for those elements that contribute to coordination and, thus, have a special function in a scientific framework. However, Stump's view hardly acknowledges components other than domain-relative theoretical principles. In comparison to Friedman, he extends the label of 'constitutive' to further theoretical principles. Yet, he does not do so in virtue of their involvement in a form of coordination, but rather because these were once empirical generalisations hardened into highly entrenched statements. In contrast, I argued that other kinds of components can be constitutive in addition to certain domain-relative theoretical principles. More importantly, it is the gradualist interpretation of the constitutive character together with the multi-layered view of coordination that justify the extension of the label of constitutive to further components, and not their degree of entrenchment in a Quinean fashion.

Entrenchment is a key word in Wimsatt's view. His perspective is based on generative entrenchment, in that certain epistemic components become more entrenched along the development of science because they generate further downstream features. Therefore, his perspective is a gradualist one, since he rejects a sharp universal separation between *a priori* and *a posteriori*, and is also a pluralist one, because he allows for several kinds of epistemic components to become generatively entrenched. However, he sees his view as strongly naturalistic, in the sense that its gradualism necessarily leads to an outright rejection of the *a priori*. However, in my view, rejecting a global divide between *a priori* and *a posteriori* does not have the same consequence, even though, as we saw, the category of the *a priori* needs thorough rethinking. Rather, the result of this rejection is viewing the attribution of a constitutive character to a component within a framework as a matter of degree, context, and comparison, and yet as a meaningful distinction to understand how justification and coordination work. Finally, my notion of 'generative potential' is clearly inspired by Wimsatt's 'generative entrenchment'.[167] Yet, I abstract away from the analogy between developmental systems and scientific development that characterises his notion. In fact, although this analogy is very powerful with respect to several examples, it obscures those cases in which the 'generativity' of some epistemic components does not depend on the historical accumulation of conceptual dependencies, but on epistemic operations that are typically human: abstraction, generalisation, idealisation, formalisation.

Lastly, my perspective is much indebted to Chang's work. His pragmatist view of metaphysical or transcendental principles underlying epistemic activities, and, more generally, his methodical analysis of the presuppositions in scientific inquiries by looking at history and practice have been a great source of inspiration. Yet, my view of constitutive elements and coordination tries to go beyond Chang's assumption that the *a priori* can be reinterpreted in terms of a single kind of general presuppositions fundamentally tied to epistemic activities. In other words, the main difference between my account and Chang's is that, in my view, qualitatively different kinds of constitutive components must be taken into account when we situate any epistemic activity in context. Certainly, my domain-independent principles underlying reasoning abilities have a lot in common with Chang's own principles; yet, the constitutive character of domain-relative theoretical components and of material components cannot be neglected when analysing coordination, even when just focusing on the basic epistemic activities that it involves.

[167] Cf. above, chapter 2, section 2.3.2.

III. The broader perspective

My view that there are qualitatively different kinds of epistemic components and that these can be constitutive to various degrees may have consequences on other debates in the philosophy of science. However, the limited scope of this dissertation only allows me to briefly hint at the impact on the bigger picture with respect to the issue of justification in science.

Recall the image of the variable landscape that I described towards the end of section 2.3.2, in contrast with Quine's web. In that section, I mentioned that philosophers such as Wimsatt and Chang emphasise how the image of a 'flat' web is far from being representative of how scientific knowledge is structured, even if we give in to idealisations. According to many contemporary perspectives on science, despite a significant degree of interconnectedness and integration, scientific inquiry sometimes results in accounts that are non-reducible to – and are overlapping with – one another, whereas some connections between them can be patchy, discontinuous, often hierarchical, and non-homogeneous in places.[168]

Rather than a flat web, scientific knowledge resembles more to an irregular landscape with elevated peaks of knowledge, vast connecting plateaux, and quicksand where knowledge is fragmentary, magmatic, and in flux. At different locations and altitudes, certain supporting structures, something like built-in beams, ensure the relative stability of the overall configuration. At the same time, this landscape is subject to constant development and change not only on its surface, but sometimes even in its deepest regions. This landscape changes as a result of the development of science, that is, as a result of our enduring effort to represent and interact with the world scientifically. This interaction is a continuous process with periods of fast transformations in some areas, and long intervals of stability or even stagnation in others.

The landscape metaphor seems to immediately point at a somehow *natural* or untainted character of scientific knowledge. This is connected with the intuition that science produces, or ought to produce, representations of the world that must depend – as much as possible – only on how the world is, whereas they should be fully independent from the interactive process that brought them about. Yet, it is crucial not to forget that this landscape is human made and, in some sense, irreducibly so. The perennial landscape-building that is the scientific interaction between people, the world, and the knowledge landscape itself is one in which goals, choices, values, and social interactions are part and

[168] Cf. Dupré's (1995) and Cartwright's (1999) by now classic views on the disunity of the sciences, and Galison & Stump (1996) for an overview on the topic. Cf. Kellert et al. (2006) for an overview of pluralist trends in philosophy of science.

parcel of the process. Along this process, at times, it becomes necessary for the human actors to use some tools, those supporting beams that I mentioned above, to jump from one peak to another, whereas, at other times, they must replace those beams that they had placed deep down in some regions as supporting structures. These are the constitutive components that I have discussed in this dissertation.

To finally get out of my metaphor, what I want to emphasise is that these constitutive components do not change just as a *natural* result of the growth of scientific knowledge. The need to replace the existing ones or placing new ones certainly comes from the constraints that the world forces us to follow in order to proceed with our epistemic interaction with it. Yet, there are often possible alternatives on where and how to place these components and, especially, what function they must have. How epistemic subjects make choices with respect to that depends on several factors, including the further empirical input and background commitments they can rely on, but also pragmatic considerations, cognitive constraints, contingent social and historical factors, and the availability of plausible alternatives. All these factors make up the justification by which epistemic subjects take one component as a supporting beam rather than another, in order to coordinate one region of the landscape at a certain altitude with its concrete world counterpart.

Acknowledgements

Personal encounters can change the course of events. This dissertation has been deeply shaped by some such encounters, for which I will always be thankful:

Maria Kronfeldner and I met in September 2015, shortly after I moved to Budapest. From the moment I started working with her, Maria has been a source of inspiration, and not just as a professional philosopher and teacher. She has been a guide in the journey through the PhD woods, a journey that has also led me to a personal rediscovery of creative and political urges. Her role as a PhD supervisor has influenced this dissertation on many levels.

Flavia Padovani showed me the human face of professional philosophy at the first academic conference I went to as an awkward audience member in January 2016. Ever since, she's been a thoughtful associate supervisor, giving me heartfelt feedback and support.

Anna Kocsis, Matthew Baxendale, Garrett Mindt, and I met in September 2015 at CEU. Anna has been an invaluable presence throughout these years thanks to her emotional and intellectual closeness. The "flying tigers" sessions with Matthew and Garrett have provided me with great feedback, a sound working routine, and, most importantly, countless good memories in Budapest. Several other PhD colleagues as well as faculty members at the Philosophy Department gave me useful feedback during various work-in-progress sessions and in-house conferences.

Mathieu Charbonneau was the instructor of a course I took in winter 2016 at CEU. During these years, Mathieu has given me a lot of guidance, commented on several chapter drafts, and organised the sessions of the CEU Science Studies Lab at which I presented my work.

In 2017/18, I spent ten months visiting the Munich Center for Mathematical Philosophy in Munich as a DAAD fellow. In Munich, I met Matteo De Benedetto and Flavia Crisciotti, with whom I have shared numerous academic pains and joys ever since, sustaining each other in our respective work. Discussions at the MCMP with Erik Curiel, Sam Fletcher, Neil Dewar, Alexander Reutlinger, and Stephan Hartmann have provided me with a lot of inspiration for this dissertation. I also thank all participants of the reading group in general philosophy of science, which I organised during the winter semester.

Stefano Canali and Eden Smith have been amazing mates at all conferences we participated together on both sides of the Atlantic. Our time spent in conversations and wandering across unknown cities has not just been stress-relieving, but an unexpected gift.

Meetings with my family and lifelong friends during these years have had an unreplaceable cheering effect. I especially thank Bianca, who has also given me formidable help with the graphics in the dissertation.

Tim has witnessed the development of this work through all its phases. He has travelled across half Europe to keep my spirits up and my brain active. By helping me see through mental walls and conceptual fogs, he contributed to massively improving the form and content of this dissertation.

LIST OF REFERENCES

Abrams, M. (2006). Infinite populations and counterfactual frequencies in evolutionary theory. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 37(2): 256-268.

Ackermann, R. J. (1985). *Data, Instruments, and Theory: A Dialectical Approach to Understanding Science*. Princeton, NJ: Princeton University Press.

Ampère, A. M. (1820). Mémoire présenté à l'Académie royale des sciences, le 2 octobre 1820, où se trouve compris le résumé de ce qui avait été lu à la même Académie les 18 et 25 septembre 1820, sur les effets des courants électriques. *Annales de Chimie et de Physique*, 15: 59-76.

Ankeny, R. A., & Leonelli, S. (2016). Repertoires: A post-Kuhnian perspective on scientific change and collaborative research. *Studies in History and Philosophy of Science Part A*, 60: 18-28.

Aoki, K. (1986). A stochastic model of gene-culture coevolution suggested by the "culture historical hypothesis" for the evolution of adult lactose absorption in humans. *Proceedings of the National Academy of Sciences*, 83(9): 2929–2933.

Appleyard, R. (1930). *Pioneers of Electrical Communication*. London: MacMillan and Co.

Archibald, T. (1988). Tension and potential from Ohm to Kirchhoff. *Centaurus*, 31(2): 141-163.

Atherton, T. (1986). A history of Ohm's law. *Electronics and Power*, 32(6): 467-472.

Auyang, S. Y. (1999). *Foundations of Complex-system Theories: in Economics, Evolutionary Biology, and Statistical Physics*. Cambridge: Cambridge University Press.

Avital, E., & Jablonka, E. (2000). *Animal Traditions: Behavioural Inheritance in Evolution*. Cambridge: Cambridge University Press.

Baird, D. (2003). Thing Knowledge: Outline of a materialist theory of knowledge. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 39-67.

—— (2004). *Thing Knowledge: A Philosophy of Scientific Instruments*. Berkeley, CA: University of California Press.

Barberousse, A., & Samadi, S. (2015). Formalising Evolutionary Theory. In: Heams, T., Huneman, P., Lecointre, G., & Silberstein, M., (Eds.), *Handbook of Evolutionary Thinking in the Sciences*. Dordrecht, the Netherlands: Springer, 229-246.

Barwich, A. S., & Chang, H. (2015). Sensory measurements: coordination and standardization. *Biological Theory*, 10(3): 200-211.

Bateson, W. (1909). *Mendel's Principles of Heredity*. Cambridge, London, New York: Cambridge University Press.

Batterman, R. W. (2001). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York and Oxford: Oxford University Press.

Beebee, H., Hitchcock, C., & Price, H., (Eds.)., (2017). *Making a Difference: Essays on the Philosophy of Causation*. Oxford: Oxford University Press.

Bitbol, M., Kerszberg, P., & Petitot, J., (Eds.), (2009). *Constituting Objectivity: Transcendental Perspectives on Modern Physics*. Dordrecht, the Netherlands: Springer.

Block, N. (1983). Mental pictures and cognitive science. *The Philosophical Review*, 92(4): 499-541.

Boon, M. (2004). Technological instruments in scientific experimentation. *International Studies in the Philosophy of Science,* 18(2-3): 221–230.

—— (2015). The scientific use of technological instruments. In: Hansson, S. O., (Ed.), *The Role of Technology in Science: Philosophical Perspectives.* Dordrecht, the Netherlands: Springer, 55-79.

Bosco, F., Castro, D., & Briones, M. R. (2012). Neutral and stable equilibria of genetic systems and the Hardy–Weinberg principle: limitations of the chi-square test and advantages of auto-correlation functions of allele frequencies. *Frontiers in Genetics*, 3: 1–10.

Boumans, M. (2005). Measurement outside the laboratory. *Philosophy of Science*, 72(5): 850-863.

—— (2007). Invariance and calibration. In: Boumans, M., (Ed.), *Measurement in Economics: A Handbook*. Amsterdam: Elsevier, 231-248.

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Bowler, P. J. (1989). *Evolution: The History of an Idea*. Berkeley, Los Angeles, London: University of California Press.

Brittan, G. G. (1978). *Kant's Theory of Science*. Princeton, NJ: Princeton University Press.

Brown, J. R. (1991). Thought experiments: A Platonic account. In: Horowitz. T., and Massey, G, (Eds.), *Thought Experiments in Science and Philosophy*. Savage, MD: Rowman & Littlefield, 119-128.

—— (2004). Peeking into Plato's heaven. *Philosophy of Science,* 71(5): 1126-1138.

Brown, T. M. (1969). The electric current in early nineteenth-century French physics. *Historical Studies in the Physical Sciences*, 1: 61-103.

Buss, L. W. (1987). *The Evolution of Individuality*. Princeton, NJ: Princeton University Press.

Buzzoni, M. (2008). *Thought Experiment in the Natural Sciences. An Operational and Reflexive-Transcendental Conception*. Würzburg: Königshausen+Neumann.

—— (2013a). Thought experiments from a Kantian point of view. In: Frappier, M., Meynell, L., & Brown, J. R., (Eds.), *Thought Experiments in Philosophy, Science, and the Arts*. New York and London: Routledge, 90-106.

—— (2013b). On thought experiments and the Kantian a priori in the natural sciences: A reply to Yiftach J. H. Fehige. *Epistemologia*, 36: 277–293.

—— (2018). Kantian accounts of thought experiments. In: Brown, J. R., Fehige, Y., Stuart, M., (Eds.), *The Routledge Companion to Thought Experiments*. New York: Routledge, 327-341.

Callebaut, W. (1993). *Taking the Naturalistic Turn, or How Real Philosophy of Science is Done*. Chicago, IL: University of Chicago Press.

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, 67(6): 380-400.

—— (1974). Evolutionary Epistemology. In: Schlipp, P. A., (Ed.), *The Philosophy of Karl Popper*. LaSalle, IL: Open Court, 413–63.

Caneva, K. L. (1978). From Galvanism to electrodynamics: the transformation of German physics and its social context. *Historical Studies in the Physical Sciences*, 9: 63-159.

—— (1980). Ampère, the etherians, and the Oersted connexion. *The British Journal for the History of Science*, 13(2): 121-138.

Carey, S. (2009). *The Origin of Concepts*. New York and Oxford: Oxford University Press.

Carnap, R. (1936). Testability and meaning. *Philosophy of Science*, 3(4): 419-471.

—— (1950). Empiricism, semantics, and ontology. *Revue Internationale de Philosophie*: 20-40.

—— (1966). *An Introduction to the Philosophy of Science*. London and New York: Basic Books.

Carus, A. W. (2007). *Carnap and Twentieth-Century Thought: Explication as Enlightenment.* Cambridge: Cambridge University Press.

Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Cartwright, N., & Bradburn, N. (2011) A theory of measurement. In: *The Importance of Common Metrics for Advancing Social Science Theory and Research: A Workshop Summary*. Washington: National Academies Press, 53-56.

Cassirer, E. (1923). *Substance and Function and Einstein's Theory of Relativity*. Chicago, IL: Open Court.

Castellani, E., (Ed.), (1998). *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*. Princeton, NJ: Princeton University Press.

Castle, W. E. (1903). The laws of heredity of Galton and Mendel, and some laws governing race improvement by selection. *Proceedings of the American Academy of Arts and Sciences*, 39: 223–242.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Chalmers, A. (2003). The theory-dependence of the use of instruments in science. *Philosophy of Science*, 70(3): 493-509.

Chang, H. (1995). Circularity and Reliability in Measurement. *Perspectives on Science* 3(2): 153-172

—— (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.

—— (2008). Contingent transcendental arguments for metaphysical principles. *Royal Institute of Philosophy Supplement* 63: 113-133.

—— (2009). Ontological principles and the intelligibility of epistemic activities. In: De Regt, H. W., Leonelli, S., & Eigner, K., (Eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press, 64–82.

—— (2011a). The persistence of epistemic objects through scientific change, *Erkenntnis* 75(3): 413-429.

—— (2011b). The philosophical grammar of scientific practice, *International Studies in the Philosophy of Science*, 25(3): 205-221.

—— (2012). *Is Water H2O? Evidence, Realism and Pluralism*. Dordrecht, the Netherlands: Springer.

—— (2014) Epistemic activities and systems of practice: Units of analysis in philosophy of science after the practice turn. In: Soler, L., Zwart, S., Lynch, M., & Israel-Jost, V., (Eds.), *Science after the Practice Turn in the Philosophy, History, and Social Studies of Science.* New York, NY: Routledge, 67-79.

—— (2015) The rise of chemical kinds through epistemic iteration. In: Kendig, C., (Ed.), *Natural Kinds and Classification in Scientific Practice*. London and New York: Routledge, 33-46.

—— (2017). Pragmatist coherence as the source of truth and reality. *Proceedings of the Aristotelian Society*, 117(2): 103-122.

—— (2020). Pragmatism, perspectivism, and the historicity of science. In: Massimi, M., & McCoy, C. D., (Eds.), *Understanding Perspectivism: Scientific Challenges and Methodological Prospects*. New York and London: Routledge, pp. 10-27.

Charbonneau, M. (2013). The cognitive life of mechanical molecular models. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4): 585-594.

—— (2014). Populations without reproduction. *Philosophy of Science*, 81(5): 727-740.

Contessa, G. (2007). Scientific representation, interpretation, and surrogative reasoning. *Philosophy of Science*, 74(1): 48-68.

Crombie, A. C. (1994). *Styles of Scientific Thinking in the European Tradition: The History of Argument and Explanation Especially in the Mathematical and Biomedical Sciences and Arts*. London: Duckworth.

Crow, J. F. (1988). Eighty years ago: the beginnings of population genetics. *Genetics*, 119(3): 473–476.

—— (1999). Hardy, Weinberg and language impediments. *Genetics*, 152(3): 821–825.

Curiel, E. (2018). Why rigid designation cannot stand on scientific ground. Manuscript submitted for publication. Available online at http://strangebeautiful.com/papers/curiel-against-rigidity.pdf. Last accessed: December 4th, 2019.

—— (2019). Schematizing the observer and the epistemic content of theories. Forthcoming in: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.

Danchin, É., Charmantier, A., Champagne, F. A., Mesoudi, A., Pujol, B., & Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews Genetics*, 12(7): 475-486.

Darwin, C. (1859). T*he Origin of Species*. London: John Murray.

—— (1881). *The Formation of Vegetable Mould, through the Action of Worms, with Observation on their Habits*. London: John Murray.

Diaconis, P. (2002). G. H. Hardy and probability??? *Bulletin of the London Mathematical Society*, 34(4): 385–402.

DiSalle, R. (2006). *Understanding Space-Time: The Philosophical Development of Physics from Newton to Einstein*. Cambridge: Cambridge University Press.

Dobzhansky, T. (1998). The birth of the genetic theory of evolution in the Soviet Union in the 1920's. In: Mayr, E., & Provine, W. B., (Eds.), *The Evolutionary Synthesis: Perspectives on the Unification of Biology.* Cambridge, MA, and London: Harvard University Press, 229–242.

Duhem, P. (1914/1954). *The Aim and Structure of Physical Theory*, 2nd edn, Princeton, NJ: Princeton University Press. Trans. Philip P. Wiener.

Dupré, J. (1995). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA, and London: Harvard University Press.

Earnshaw, E. (2015). Evolutionary forces and the Hardy-Weinberg equilibrium. *Biology and Philosophy*, 30(3): 423–437.

Eberhardt, F. (2011). Reliability via synthetic a priori: Reichenbach's doctoral thesis on probability. *Synthese*, 181(1): 125-136.

Edwards, A. W. F. (2008). G. H. Hardy (1908) and Hardy–Weinberg Equilibrium. *Genetics*, 179(3): 1143–1150.

Elgin, M. (2003). Biology and a priori laws. *Philosophy of Science*, 70(5): 1380–1389.

Erwin, D. H. (2008). Macroevolution of ecosystem engineering, niche construction and diversity. *Trends in Ecology & Evolution*, 23(6): 304-310.

Everett, J. (2015). The constitutive a priori and the distinction between mathematical and physical possibility. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 52: 139-152.

Feest, U. (2010). Concepts as tools in the experimental generation of knowledge in cognitive neuropsychology. *Spontaneous Generations: A Journal for the History and Philosophy of Science*, 4(1): 173-190.

Feest, U., & Steinle, F., (Eds.), (2012). *Scientific Concepts and Investigative Practice*. Berlin: de Gruyter.

Fehige, Y. (2012). 'Experiments of pure reason': Kantianism and thought experiments in science. *Epistemologia,* 35: 141-160.

—— (2013). The relativized a priori and the laboratory of the mind: Towards a neo-Kantian account of thought experiments in science. *Epistemologia*, 36: 55-73.

Fehige, Y., & Stuart, M. T. (2014). On the origins of the philosophy of thought experiments: The forerun. *Perspectives on Science*, 22(2): 179-220.

Feldman, M. W., & Cavalli-Sforza, L. L. (1989). On the theory of evolution under genetic and cultural transmission with application to the lactose absorption problem. In: Feldman, M. W., (Ed.), *Mathematical Evolutionary Theory*. Princeton, NJ: Princeton University Press, 145–173.

Ferrari, M. (2012). Between Cassirer and Kuhn. Some remarks on Friedman's relativized a priori. *Studies in History and Philosophy of Science Part A*, 43(1): 18-26.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52: 399–433.

—— (1922). On the dominance ratio. *Proceedings of the Royal Society of Edinburgh*, 42: 321–341.

—— (1930/1999) *The Genetical Theory of Natural Selection.* Oxford: Oxford University Press.

Fletcher, S. (2017). The principle of stability. Forthcoming in: *Philosophers' Imprint.*

Franklin, L. R. (2005). Exploratory experiments. *Philosophy of Science*, 72(5): 888-899.

Friedman, M. (1992). *Kant and the Exact Sciences.* Cambridge, MA: Harvard University Press.

—— (1993). Remarks on the history of science and the history of philosophy. In: Horwich, P., (Ed.), *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, MA: MIT Press, 36-54.

—— (1994). Geometry, convention, and the relativized a priori: Reichenbach, Schlick, and Carnap. In: Salmon, W., & Wolters, G., (Eds.), *Logic, Language, and the Structure of Scientific Theories*. Pittsburgh/Konstanz: University of Pittsburgh Press/Universitätsverlag Konstanz, 21-34.

—— (1999). *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.

—— (2000a). *A Parting of the Ways: Carnap, Cassirer, and Heidegger*. Chicago, IL: Open Court.

—— (2000b). Transcendental philosophy and a priori knowledge: A neo-Kantian perspective. In: Boghossian, P., & Peacocke, C., (Eds.), *New Essays on the A Priori*. Oxford: Clarendon Press, 367-383.

—— (2001). *Dynamics of Reason*. Stanford, CA: Csli Publications.

—— (2009). Einstein, Kant, and the relativized *a priori*. In: Bitbol, M., Kerszberg, P., & Petitot, J., (Eds.), *Constituting Objectivity: Transcendental Perspectives on Modern Physics*. Dordrecht, the Netherlands: Springer, 253-267.

—— (2010). Synthetic history reconsidered. In: Domski, M., & Dickson, M., (Eds.), *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science*. Chicago and La Salle, IL: Open Court, 569-813.

—— (2012). Reconsidering the dynamics of reason: Response to Ferrari, Mormann, Nordmann, and Uebel. *Studies in History and Philosophy of Science Part A*, 43(1): 47-53.

Frigerio, A., Giordani, A., & Mari, L. (2010). Outline of a general model of measurement. *Synthese*, 175(2): 123-149.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521): 1211-1221.

Galison, P. (1987). *How Experiments End*. Chicago, IL: University of Chicago Press.

—— (1988). History, philosophy, and the central metaphor. *Science in Context*, 2(1): 197-212.

Galison, P. L., & Stump, D. J., (Eds.), (1996). *The Disunity of Science: Boundaries, Contexts, and Power*. Stanford, CA: Stanford University Press.

Gerbault, P., Liebert, A., Itan, Y., Powell A., Currat, M., Burger, J., Swallow, D. M., & Thomas, M. G. (2011). Evolution of Lactase Persistence: An Example of Human Niche Construction. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 366(1566): 863–77.

Germain, P. L. (2014). Living instruments and theoretical terms: Xenografts as measurements in cancer research. In: Galavotti, M. C., Dieks, D., Gonzalez, W. J., Hartmann, S., Uebel, T., & Weber, M. (Eds.). *New Directions in the Philosophy of Science*. Dordrecht, the Netherlands: Springer, 141-155.

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

—— (2010). *Scientific Perspectivism*. Chicago, IL: University of Chicago Press.

Gillespie, J. H. (1998). *Population Genetics: A Concise Guide*. Baltimore, MD, and London: The Johns Hopkins University Press.

Gillmor, C. S. (1972). *Coulomb and the Evolution of Physics and Engineering in Eighteenth-Century France.* Princeton, NJ: Princeton University Press

Giordani, A., & Mari, L. (2012). Measurement, models, and uncertainty. *IEEE Transactions on Instrumentation and Measurement*, 61(8): 2144–2152.

Godfrey-Smith, P. (2009a). Abstractions, idealizations, and evolutionary biology. In: Barberousse, A., Morange, M., & Pradeu, T., (Eds.), *Mapping the Future of Biology: Concepts and Theories.* Dordrecht, the Netherlands: Springer, 47-56.

—— (2009b). *Darwinian Populations and Natural Selection*. Oxford and New York: Oxford University Press.

Gooding, D. (1990). *Experiment and the Making of Meaning: Human Agency in Scientific Observation and Experiment.* Dordrecht, Boston, and London: Kluwier.

—— (2003). Varying the cognitive span: Experimentation, visualization, and computation. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 255-284.

Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, IN: Hackett publishing.

Gouvea, D. J. (2015). Explanation and the evolutionary first law(s). *Philosophy of Science*, 82(3): 363–382.

Griesemer, J. R. (1990). Material models in biology. In *PSA: Proceedings of the Biennial meeting of the Philosophy of Science Association*, Vol. 1990, No. 2, 79-93.

—— (2013). Formalization and the meaning of "theory" in the inexact biological sciences. *Biological Theory*, 7(4): 298-310.

Griffiths, P. E., & Gray, R. D. (1994). Developmental systems and evolutionary explanation. *The Journal of Philosophy,* 91(6): 277–304.

Gupta, M. S. (1980). Georg Simon Ohm and Ohm's Law. *IEEE Transactions on Education*, 23(3): 156-162.

Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science.* Cambridge: Cambridge University Press.

—— (1992). 'Style' for historians and philosophers. *Studies in History and Philosophy of Science Part A*, 23(1): 1-20.

—— (1993). Working in a new world: The taxonomic solution. In: Horwich, P., (Ed.), *World Changes*. Cambridge, MA: MIT Press, 275-311.

—— (1996). The disunities of the sciences. In: Galison, P. L. & Stump, D. J., (Eds.), *The Disunity of Science: Boundaries, Contexts, and Power.* Stanford, CA: Stanford University Press, 37-74.

—— (2015). Let's not talk about objectivity. In: Padovani, F., Richardson, A., & Tsou, J. Y., (Eds.), *Objectivity in Science: New Perspectives from Science and Technology Studies*. Cham: Springer, 19-33.

Haldane, J. B. S. (1924). Part I. A mathematical theory of natural and artificial selection. *Transactions of the Cambridge Philosophical Society*, 23: 19–41.

Hamilton, W. D. (1963). The Evolution of Altruistic Behavior. *American Naturalist,* 97(896): 354-56.

Hanson, N. R. (1958). *Patterns of Discovery*. Cambridge: Cambridge University Press.

Hardy, G. H. (1908). Mendelian proportions in a mixed population. *Science*, 28: 49–50.

Harré, R. (2003). The materiality of instruments in a metaphysics for experiments. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 19-38.

Hartl, D. L., & Clark, A. G. (2007). *Principles of Population Genetics* (4th ed.). Sunderland, MA: Sinauer Associates.

Heidelberger, M. (1980). Towards a logical reconstruction of revolutionary change: the case of Ohm as an example. *Studies in History and Philosophy of Science Part A*, 11(2): 103-121.

—— (2003). Theory-ladenness and scientific instruments in experimentation. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 138-151.

Heilbron, J. L. (1979). *Electricity in the 17th and 18th Century. A Study of Early Modern Physics.* Berkeley, CA: University of California Press.

Helmholtz, H. L. F. von (1867). *Handbuch der physiologischen Optik*. Leipzig: L. Voss.

Hintikka, J. (1959). Existential presuppositions and existential commitments. *The Journal of Philosophy*, 56(3): 125-137.

Holden, C., & Mace, R. (1997). Phylogenetic analysis of the evolution of lactose digestion in adults. *Human Biology*, 69(5): 605-628.

Holmes, P., & Shea-Brown, E. T. (2006). Stability. *Scholarpedia*, 1(10): 1838.

Howard, D. (2010). "Let me briefly indicate why I do not find this standpoint natural." Einstein, General Relativity, and the Contingent A Priori. In: Domski, M., & Dickson, M., (Eds.), *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science.* Chicago and La Salle: Open Court, 333-355.

Hull, D. L. (1988). *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago, IL: University of Chicago Press.

Jablonka, E., & Lamb, M. J. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioural and Symbolic Variation in the History of Life*. Cambridge, MA: MIT Press.

Johansson, I. (2014). Constancy and circularity in the SI. *Metrology Bytes*. www.metrologybytes.net/PapersUnpub/OpEds/Johansson_2014.pdf. Last accessed: 2 February 2018.

Jongeling, T. B. (1985). On an axiomatization of evolutionary theory. *Journal of Theoretical Biology*, 117(4): 529-543.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2): 136-153.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In: Kahneman, D., Slovic, P., & Tversky, A., (Eds.), *Judgement Under Uncertainty: Heuristics and Biases*. New York, NY: Cambridge University Press, 201–11.

Kellert, S. H., Longino, H. E., & Waters, C. K., (Eds.), (2006). *Scientific Pluralism*. Minneapolis, MN: University of Minnesota Press.

Kendig, K., (Ed.), (2015). *Natural Kinds and Classification in Scientific Practice*. London and New York: Routledge.

Kosso, P. (1989). *Observability and Observation in Physical Science*. Dordrecht, the Netherlands: Kluwer.

Kroes, P. (2003). Physics, experiments, and the concept of nature. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 68-86.

Kuhn, T. S. (1962/1970). The structure of scientific revolutions. *International Encyclopedia of Unified Science*, vol. 2, no. 2.

—— (1970). Second thoughts on paradigms. In: Suppe, F., (Ed.), (1974), *The Structure of Scientific Theories*. Urbana, IL: University of Illinois Press, 459-482.

—— (1990). Dubbing and redubbing: the vulnerability of rigid designation. In: Savage, C. W., (Ed.), *Scientific Theories*. Minneapolis, MN: University of Minnesota Press, 298-318.

—— (2000). The road since *Structure*. In: Conant, J., & Haugeland, T., (Eds.), *The Road since Structure. Philosophical essays, 1970-1993, with an Autobiographical Interview*. Chicago and London: The University of Chicago Press, 90-103.

Laland, K. N., Odling-Smee, F. J., & Feldman, M. W. (1996). The evolutionary consequences of niche construction: a theoretical investigation using two-locus theory. *Journal of Evolutionary Biology*, 9(3): 293-316.

—— (1999). Evolutionary consequences of niche construction and their implications for ecology. *Proceedings of the National Academy of Sciences*, 96(18): 10242-10247.

Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: bringing genetics and the human sciences together. *Nature Reviews Genetics*, 11(2): 137.

Laland, K., Uller, T., Feldman, M., Sterelny, K., Müller, G. B., Moczek, A., Jablonka, E., Wray, G. A., Hoekstra, H. E., Futuyma, D. J., Lenski, R. E., Mackay, T. F. C., Schluter, D., & Strassman, J. E. (2014). Does evolutionary theory need a rethink? *Nature News*, 514(7521): 161-164.

Lakatos, I. (1970). Falsification and the methodology of scientific research programs. In: Lakatos, I., & Musgrave, A., (Eds.), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 91-196.

—— (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: Cambridge University Press.

Leonelli, S. (2007). Growing weed, producing knowledge: an epistemic history of Arabidopsis thaliana. *History and Philosophy of the Life Sciences*, 193-223.

—— (2009). Understanding in biology: the impure nature of biological knowledge. In: De Regt, H. W., Leonelli, S., & Eigner, K., (Eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press, 189-209.

Lewis, C. I. (1929). *Mind and the World-Order: Outline of a Theory of Knowledge*. New York, NY: Dover Publications.

Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1(1): 1-18.

—— (1982). Organism and environment. In: Plotkin, H. C., (Ed.)., *Learning, Development and Culture*. New York, NY: Wiley, 151–170.

—— (1983). Gene, organism, and environment. In: Bendall, D. S., (Ed.), (1985). *Evolution from Molecules to Men*. New York, NY: Cambridge University Press, 273-285.

—— (1985). Adaptation. In: Levins, R., & Lewontin, R. C., (Eds.), *The Dialectical Biologist*. Cambridge, MA: Harvard University Press, 65-84.

Li, C. C. (1988). Pseudo-random mating populations. In celebration of the 80th anniversary of the Hardy-Weinberg law. *Genetics*, 119(3): 731–737.

Love, A. C. (2013). Theory is as theory does: scientific practice and theory structure in biology. *Biological Theory*, 7(4): 325-337.

—— (Ed.), (2015). *Conceptual Change in Biology*. New York, NY: Springer.

Mach, E. (1896/1986). *Principles of the Theory of Heat*. T.J. McCormack (trans.), Dordrecht, the Netherlands: D. Reidel.

Mari, L. (2003). Epistemology of measurement. *Measurement*, 34(1): 17-30.

Massimi, M. (2008). Why there are no ready-made phenomena: What philosophers of science should learn from Kant. *Royal Institute of Philosophy Supplements*, 63: 1-35.

Masterman, M. (1970). The nature of a paradigm. In: Lakatos, I., & Musgrave, A., (Eds.), *Criticism and the Growth of Knowledge*. London: Cambridge University Press, 59-89.

Matthen, M., & Ariew, A. (2002). Two ways of thinking about fitness and natural selection. *Journal of Philosophy*, 99(2): 55–83.

—— (2009). Selection and causation. *Philosophy of Science*, 76(2): 201–24.

Maynard Smith, J. M. (1976). Group selection. *Quarterly Review of Biology*, 51(2): 277–83.

Maynard Smith, J. M., & Szathmary, E. (1995). *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Mayo, O. (2008). A century of Hardy–Weinberg equilibrium. *Twin Research and Human Genetics*, 11(3): 249–256.

Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Harvard University Press.

Mayr, E., & Provine, W. B. (Eds.), (1998). *The Evolutionary Synthesis: Perspectives on the Unification of Biology*. Cambridge, MA: Harvard University Press.

McClimans, L. (2013). The role of measurement in establishing evidence. *Journal of Medicine and Philosophy*, 38(5): 520-538.

McClimans, L., Browne, J., & Cano, S. (2017). Clinical outcome measurement: Models, theory, psychometrics and practice. *Studies in History and Philosophy of Science Part A*, 65: 67-73.

McKnight, J. L. (1967). Laboratory notebooks of GS Ohm: a case study in experimental method. *American Journal of Physics*, 35(2): 110-114.

McShea, D. W., & Brandon, R. N. (2010). *Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems*. Chicago, IL: University of Chicago Press.

Merlin, F. (2017). Limited extended inheritance. In: Huneman, P., & Walsh, D., (Eds.), *Challenges to Evolutionary Theory. Development, Inheritance*, *Adaptation*. Oxford: Oxford University Press, 263-279.

Mets, A. (2019). A philosophical critique of the distinction of representational and pragmatic measurements on the example of the periodic system of chemical elements. *Foundations of Science*, 24(1): 73-93.

Michel, M. (2019). The mismeasure of consciousness: A problem of coordination for the Perceptual Awareness Scale. *Philosophy of Science*, 86(5): 1239-1249.

Michod, R. E. (2000). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton, NJ: Princeton University Press.

Millstein, R. L., Skipper Jr, R. A., & Dietrich, M. R. (2009). (Mis) interpreting mathematical models: drift as a physical process. *Philosophy & Theory in Biology*, 1: 1-13.

Morgan, M. S. (2003). Experiments without material intervention: Model experiments, virtual experiments and virtually experiments. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 216-235.

Mormann, T. (2012). A place for pragmatism in the dynamics of reason? *Studies in History and Philosophy of Science Part A*, 43(1): 27–37.

Morrison, M. (2007). The development of population genetics. In: Matthen, M., & Stephens, C., (Eds.), *Philosophy of Biology*. Amsterdam, the Netherlands: Elsevier, 309–333.

—— (2011). One phenomenon, many models: Inconsistency and complementarity. *Studies in History and Philosophy of Science Part A*, 42(2): 342–351.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. *Cognitive Models of Science*, 15: 3-44.

—— (2010). *Creating Scientific Concepts*. Cambridge, MA: MIT press.

Nikolov, S., Yankulova, E., Wolkenhauer, O., & Petrov, V. (2007). Principal difference between stability and structural stability (robustness) as used in systems biology. *Nonlinear Dynamics, Psychology, and Life Sciences*, 11(4): 413-33.

Norton, J. D. (1991). Thought experiments in Einstein's work. In: Horowitz. T., & Massey, G., (Eds.), *Thought Experiments in Science and Philosophy*. Savage, MD: Rowman & Littlefield, 129-148.

—— (2004). Why thought experiments do not transcend empiricism. In: Hitchcock, C. R., (Ed.), *Contemporary Debates in Philosophy of Science*. Malden, Oxford, and Carlton: Blackwell, 44-66.

—— (2010). There are no universal rules for induction. *Philosophy of Science*, 77(5): 765–777.

—— (2014). A material dissolution of the problem of induction. *Synthese*, 191(4): 671–690.

Odling-Smee F. J. (1988). Niche constructing phenotypes. In: Plotkin, H. C., (Ed.)., *The Role of Behavior in Evolution*, 73–132. Cambridge, MA: MIT Press.

—— (2007). Niche inheritance. A possible basis for classifying multiple inheritance systems in evolution. *Biological Theory,* 2(3): 276-289.

Odling-Smee, F. J. & Laland, K. N. (2011). Ecological inheritance and cultural inheritance: What are they and how do they differ? *Biological Theory,* 6(3): 220-230.

Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (1996). Niche construction. *American Naturalist,* 147(4): 641–648.

—— (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.

Ørsted, H. C., & Fourier, J. B. (1823). Sur quelques nouvelles expériences thermoélectriques faites par M. le Baron Fourier et M. Oersted. *Annales de Chimie et de Physique*, 22: 375-389.

Ohm, G. S. (1825a). Vorläufige Anzeige des Gesetzes, nach welchem Metalle die Contact-Elektricität leiten. *Schweigger's Journal für Chemie und Physik*, Bd. 44: 110-118, and Poggendorff´s Annalen, Bd. 4: 79-88. Reprinted in: Lommel, E., (Ed.), (1892). *Gesammelte Abhandlungen von G.S Ohm*. Leipzig: Arthur Meiner, 1-8.

—— (1825b). Ueber Leitungsfähigkeit der Metalle für Elektricität. *Schweigger's Journal für Chemie und Physik*, Bd. 44: 245-247. Reprinted in: Lommel, E., (Ed.), (1892). *Gesammelte Abhandlungen von G.S Ohm*. Leipzig: Arthur Meiner, 9-10.

—— (1825c). Ueber Elektricitätsleiter. *Schweigger's Journal für Chemie und Physik*, Bd. 44: 370-373. Reprinted in: Lommel, E., (Ed.), (1892). *Gesammelte Abhandlungen von G.S Ohm*. Leipzig: Arthur Meiner, 11-13.

—— (1826). Bestimmung des Gesetzes, nach welchem Metalle die Contactelektricität leiten, nebst einem Entwurfe zu einer Theorie des Voltaischen Apparates und des Schweiggerschen Multiplicators. *Schweigger's Journal ´für Chemie und Physik*, Bd. 46: 137-166. Reprinted in: Lommel, E., (Ed.), (1892). *Gesammelte Abhandlungen von G.S Ohm*. Leipzig: Arthur Meiner, 14-36.

—— (1827). *Die galvanische Kette, matematisch bearbeitet*. Translated by Francis, W. (1891). *The Galvanic Circuit Investigated Mathematically*. New York, NY: D. Van Nostrand Company.

Okasha, S. (2005). Multi-level selection and the major transitions in evolution. *Philosophy of Science*, 72(5): 1013-25.

—— (2018). The strategy of endogenization in evolutionary biology. *Synthese*, https://doi.org/10.1007/s11229-018-1832-6.

Olby, R. C. (1966). *Origins of Mendelism*. London: Constable & Co.

Ospovat, D. (1995). *The Development of Darwin's Theory: Natural History, Natural Theology, and Natural Selection, 1838-1859*. Cambridge: Cambridge University Press.

Padovani, F. (2011). Relativizing the relativized a priori: Reichenbach's axioms of coordination divided. *Synthese*, 181(1): 41-62.

—— (2015). Measurement, coordination, and the relativized a priori. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 52: 123-128.

—— (2017). Coordination and measurement: What we get wrong about what Reichenbach got right. In: Massimi, M., Romeijn, J. W., & Schurz, G., (Eds.), *EPSA15 Selected Papers. European Studies in Philosophy of Science*, 5, 49–60.

Pancaldi, G. (1990). Electricity and life. Volta's path to the battery. *Historical Studies in the Physical and Biological Sciences*, 21(1): 123-160.

Pap, A. (1944). The different kinds of a priori. *The Philosophical Review*, 53(5): 465-484.

Parker, W. S. (2009). Does matter really matter? Computer simulations, experiments, and materiality. *Synthese*, 169(3): 483-496.

Parrini, P. (2009). Carnap's Relativised A Priori and Ontology. In: Bitbol, M., Kerszberg, P. & Petitot, J., (Eds.), *Constituting Objectivity: Transcendental Perspectives on Modern Physics.* Dordrecht, the Netherlands: Springer, 127-143.

Pearson, K. (1904). On a generalized theory of alternative inheritance, with special reference to Mendel's laws. *Philosophical Transactions of the Royal Society*, A, 203: 53–86.

Pigliucci, M. (2007). Do we need an extended evolutionary synthesis? *Evolution: International Journal of Organic Evolution*, 61(12): 2743-2749.

Pigliucci, M., & Müller, G. B. (Eds.), (2010). *Evolution: The Extended Synthesis*. Cambridge, MA: MIT Press.

Pitts, J. B. (2018). Kant, Schlick and Friedman on space, time and gravity in light of three lessons from particle physics. *Erkenntnis*, 83(2): 135-161.

Poincaré, H. (1902). *La Science et l'Hypothèse*. English translation in: *Science and Hypothesis*. New York: Dover Books, 1952.

Polanyi, M. (1966). *The Tacit Dimension*. Chicago and London: University of Chicago Press.

Popper, K. R. (1972). *Objective Knowledge: An Evolutionary Approach.* Oxford: Clarendon Press.

Price, G. R. (1970). Selection and Covariance. *Nature,* 227(5257): 520-21.

—— (1972). Extension of Covariance Selection Mathematics. *Annals of Human Genetics,* 35(4): 485-90.

Provine, W. B. (1971). *The Origins of Theoretical Population Genetics*. Chicago and London: University of Chicago Press.

Pugh, C., & Peixoto, M. M. (2008). Structural stability. *Scholarpedia*, 3(9): 4008.

Putnam, H. (1962). The analytic and the synthetic. In: Feigl, H., & Maxwell, G., (Eds.), *Scientific Explanation, Space, and Time*. Minneapolis, MN: University of Minnesota Press. Reprinted in: Putnam, H. (1975). *Mind, Language, and Reality: Philosophical Papers*. Cambridge: Cambridge University Press, 33-69.

—— (1976). 'Two dogmas' revisited. In: Ryle, G., (Ed.), *Contemporary Aspects of Philosophy*. Stocksfield: Oriel Press. Reprinted in: Putnam, H. (1983). *Realism and Reason: Philosophical Papers*. Cambridge: Cambridge University Press, 87-97.

—— (1979). Analyticity and apriority: Beyond Wittgenstein and Quine. In: French, P. A., Uehling, T. E., & Wettstein, H. K., (Eds.), *Studies in Metaphysics*. Minneapolis, MN: University of Minnesota Press. Reprinted in: Putnam, H. (1983). *Realism and Reason: Philosophical Papers*. Cambridge: Cambridge University Press, 115-138.

Quine, W. V. O. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*: 20-43.

Radder, H. (2003). Technology and theory in experimental science. In: Radder, H., (Ed.), *The Philosophy of Scientific Experimentation*. Pittsburgh, PA: University of Pittsburgh Press, 152-173.

Reichenbach, H. (1920). *Relativitätstheorie und Erkenntnis Apriori*. Berlin: Springer. English translation in: *The Theory of Relativity and A Priori Knowledge*. Berkeley/Los Angeles, CA: University of California Press, 1965.

—— (1956). *The Direction of Time*. Berkeley, CA: University of California Press.

Rheinberger, H. J. (1997). *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube.* Stanford, CA: Stanford University Press.

—— (2013). Heredity in the twentieth century: some epistemological considerations. *Public Culture*, 30(3): 477–493.

Rheinberger, H. J., & Müller-Wille, S. (2017). *The Gene: From Genetics to Postgenomics*. Chicago, IL: University of Chicago Press.

Richardson, A. W. (2002). Narrating the history of reason itself: Friedman, Kuhn, and a constitutive a priori for the twenty-first century. *Perspectives on Science*, 10(3): 253-274.

Ruse, M. (1971). Is the theory of evolution different? Pt.1, The Central Core of the Theory. *Scientia*, 106: 765–83.

Russell, P. J. (2010). *iGenetics: A Molecular Approach* (3rd ed.). San Francisco, CA: Pearson - Benjamin Cummings.

Ruthenberg, K., & Chang, H. (2017). Acidity: Modes of characterization and quantification. *Studies in History and Philosophy of Science Part A*, 65: 121-131.

Ryckman, T. (2005). *The Reign of Relativity*. Oxford: Oxford University Press.

Schagrin, M. L. (1963). Resistance to Ohm's law. *American Journal of Physics*, 31(7): 536-547.

Schank, J. C., & Wimsatt, W. C. (1986). Generative entrenchment and evolution. In *PSA: Proceedings of the biennial meeting of the philosophy of science association*, 2: 33-60.

Scheiner, S. (2010). Toward a conceptual framework for biology. *The Quarterly Review of Biology*, 85(3): 293-318.

Shaffer, M. J. (2011). The constitutive a priori and epistemic justification. In: Shaffer, M. J., & Veber, M. L., (Eds.), *What Place for the a Priori?* Chicago, IL: Open Court Publishing, 193-209.

Shedd, J. C., & Hershey, M. D. (1913). The history of Ohm's law. *Popular Science Monthly*, 83: 599-614.

Sherry, D. (2011). Thermoscopes, thermometers, and the foundations of measurement. *Studies in History and Philosophy of Science Part A*, 42(4): 509-524.

Sober, E. (1984). *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge, MA: MIT Press.

Soler, L., Sankey, H., & Hoyningen-Huene, P., (Eds.), (2008). *Rethinking Scientific Change and Theory Comparison.* Berlin: Springer.

Stark, A. E. (2006a). Stages in the evolution of the Hardy-Weinberg law. *Genetics and Molecular Biology*, 29(4): 589–594.

—— (2006b). A clarification of the Hardy-Weinberg law. *Genetics*, 174(3): 1695–1697.

—— (2007). On extending the Hardy-Weinberg law. *Genetics and Molecular Biology*, 30(3): 664–666.

Stark, A. E., & Seneta, E. (2013). A reality check on Hardy–Weinberg. *Twin Research and Human Genetics*, 16(4): 782–789.

Steinle, F. (1997). Entering new fields: Exploratory uses of experimentation. *Philosophy of Science*, 64: S65-S74.

—— (1998). Exploratives vs. theoriebestimmtes Experimentieren: Ampères erste Arbeiten zum Elektromagnetismus. In: Heidelberger, M., & Steinle, F., (Eds.), *Experimental Essays – Versuche zum Experiment*. Baden-Baden: Nomos Verlagsgesellschaft, 272-297.

—— (2002). Experiments in history and philosophy of science. *Perspectives on Science*, 10(4). 408-432.

—— (2016). *Exploratory Experiments. Ampère, Faraday, and the Origins of Electrodynamics*. Pittsburgh, PA: University of Pittsburgh Press.

Stephens, C. (2004). Selection, drift, and the "forces" of evolution. *Philosophy of Science*, 71(4): 550–570.

Sterelny, K., Smith, K. C., & Dickinson M. (1996). The extended replicator. *Biology and Philosophy,* 11(3): 377–403.

Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, 103(2684): 677–680.

Stöltzner, M. (2009). Can the principle of least action be considered a relativized a priori? In: Bitbol, M., Kerszberg, P., & Petitot, J., (Eds.), *Constituting Objectivity: Transcendental Perspectives on Modern Physics.* Dordrecht, the Netherlands: Springer, 215-227.

Strawson, P. F. (1950). On referring. *Mind*, 59(235): 320-344.

Stuart, M. T. (2017). Imagination: A sine qua non of science. *Croatian Journal of Philosophy*, 49(17): 9–32.

Stump, D. J. (2003). Defending conventions as functionally a priori knowledge. *Philosophy of Science*, 70(5): 1149-1160.

—— (2015) *Conceptual Change and the Philosophy of Science: Alternative Interpretations of the A Priori*. New York and London: Routledge.

Suárez, M. (2012). Science, philosophy and the a priori. *Studies in History and Philosophy of Science Part A*, 43(1): 1-6.

Tal, E. (2011). How Accurate is the Standard Second? *Philosophy of Science* 78(5): 1082–96.

—— (2013). Old and new problems in philosophy of measurement. *Philosophy Compass*, 8(12): 1159–1173.

—— (2016a). How Does Measuring Generate Evidence? The Problem of Observational Grounding. In *Journal of Physics: Conference Series*, 772(1), DOI: 10.1088/1742-6596/772/1/012001.

—— (2016b). Making time: A study in the epistemology of measurement. *The British Journal for the Philosophy of Science*, 67(1): 297-335.

—— (2017a). Measurement in Science. In: Zalta, E. N.: *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), URL = <https://plato.stanford.edu/archives/fall2017/entries/measurement-science/>.

—— (2017b). A model-based epistemology of measurement. In: Mößner, N., & Nordmann, N., (Eds.), *Reasoning in Measurement*. London and New York: Routledge, 233-253.

—— (2017c). Calibration: Modelling the measurement process. *Studies in History and Philosophy of Science Part A*, 65: 33-45.

—— (2018). Naturalness and convention in the International System of Units☆. *Measurement*, 116: 631-643.

—— (2019). Individuating quantities. *Philosophical Studies*, 176(4): 853-878.

Teller, P. (2001). Twilight of the perfect model model. *Erkenntnis*, 55(3): 393-415.

—— (2004). The law-idealization. *Philosophy of Science*, 71(5): 730-741.

Tennant, N. (2014). The logical structure of evolutionary explanation and prediction: Darwinism's fundamental schema. *Biology & Philosophy*, 29(5): 611-655.

Thagard, P. (1992). *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.

—— (2012). *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change*. Cambridge, MA: MIT Press.

Thompson, P. (2007). Formalisations of evolutionary biology. In: Matthen, M., & Stephens, C., (Eds.), *Handbook of the Philosophy of Biology*. Elsevier, 485-523.

Tsou, J. Y. (2010). Putnam's account of apriority and scientific change: its historical and contemporary interest. *Synthese*, 176(3): 429–445.

—— (2012). Intervention, causal reasoning, and the neurobiology of mental disorders: Pharmacological drugs as experimental instruments. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(2): 542-551.

Uebel, T. (2012). De-synthesizing the relative a priori. *Studies in History and Philosophy of Science Part A*, 43(1): 7-17.

van Fraassen, B. C. (1968). Presupposition, implication, and self-reference. *The Journal of Philosophy*, 65(5): 136-152.

—— (1987). The semantic approach to scientific theories. In: Nersessian, N. J., (Ed.), *The Process of Science*. Dordrecht, the Netherlands: Springer, 105-124.

—— (2008). *Scientific Representation: Paradoxes of Perspective*. New York, NY: Oxford University Press.

Volta, A. (1800). On the electricity excited by the mere contact of conducting substances of different kinds, *Philosophical Transactions*, 90: 403-431.

—— (1918). *Opere*. Milano: Hoepli.

Waismann, F. (1968). Verifiability. In: Harré, R., (Ed.), *How I See Philosophy*. London: Palgrave Macmillan, 39-66.

Walsh, D. M. (2007). The pomp of superfluous causes: the interpretation of evolutionary theory. *Philosophy of Science*, 74(3): 281–303.

Weber, M. (2017). Causal selection vs causal parity in biology: relevant counterfactuals and biologically normal interventions. Forthcoming in: Waters, C. K., & Woodward, J., (Eds.), *Philosophical Perspectives on Causal Reasoning in Biology*. Minneapolis, MN: University of Minnesota Press.

Weinberg, W. (1908). Über den Nachweis der Vererbung beim Menschen. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg, Stuttgart, 64: 369-382. [On the demonstration of inheritance in humans]. Translation by Jameson, R. A., printed in Jameson, D. L. (Ed.), (1977), *Benchmark Papers in Genetics, Volume 8: Evolutionary Genetics*. Stroudsburg, PA: Dowden, Hutchinson and Ross, 115–125.

Williams, G. C. (1966). *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton, NJ: Princeton University Press.

Williams, M. (1970). Deducing the Consequences of Selection: A Mathematical Model. *Journal of Theoretical Biology*, 29(3): 343–385.

—— (1973). The Logical Status of Natural Selection and other Evolutionary Controversies: Resolution by Axiomatization. In: Bunge, M., (Ed.), *The Methodological Unity of Science.* Dordrecht, the Netherlands: D. Reidel, 84–102.

Wimsatt, W. C. (1987). Generative entrenchment, scientific change, and the analytic–synthetic distinction: A developmental model of scientific evolution. In: *Invited address at the Western (now Central) Division of the APA meetings, unpublished ms*.

—— (1999). Generativity, entrenchment, evolution, and innateness: philosophy, evolutionary biology, and conceptual foundations of science. In: Hardcastle, V. G., (Ed.), *Where Biology Meets Psychology: Philosophical Essays.* Cambridge, MA, and London: MIT Press, 137–179.

—— (2007). *Re-engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Winter, H. J. J. (1944). The reception of Ohm's electrical researches by his contemporaries. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 35(245): 371-386.

Winther, R. G. (2015). Mapping kinds in GIS and cartography. In: Kendig, C., (Ed.), *Natural Kinds and Classification in Scientific Practice.* London and New York: Routledge, 197–216.

Winther, R. G., Wade, M. J., & Dimond, C. C. (2013). Pluralism in evolutionary controversies: styles and averaging strategies in hierarchical selection theories. *Biology and Philosophy*, 28(6): 957-979.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation.* New York and Oxford: Oxford University Press.

—— (2006). Some varieties of robustness. *Journal of Economic Methodology*, 13(2): 219-240.

Wright, S. (1917). Color inheritance in mammals. 6. Cattle. *Journal of Heredity*, 8: 521–527.

—— (1918). Color inheritance in mammals. 11. Man. *Journal of Heredity*, 9: 231–232.

Yule, G. U. (1902). Mendel's laws and their probable relations to intra-racial heredity. *New Phytologist*, 1: 193–207, 222–238.