Studies in History and Philosophy of Science 106 (2024) 1–11

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

From the philosophy of measurement to the philosophy of classification: Generalizing the problem of coordination and historical coherentism

François Papale [a,b,*]

[a] Université Laval, Faculty of Philosophy, Québec, Canada
[b] Université de Sherbrooke, Department of Microbiology and Infectious Disease, Québec, Canada

ARTICLE INFO

Keywords:
Epistemology
Classification
Empiricism
Problem of coordination
Phylogenetics
Historical coherentism

ABSTRACT

The objective of this paper is twofold. First, I present a framework called historical coherentism (Chang, 2004; Tal, 2016; Van fraassen 2008) and argue that it is the best epistemological framework available to tackle the problem of coordination, an epistemic conundrum that arises with every attempt to provide empirical content to scientific theories, models or statements. Second, I argue that the problem of coordination, which has so far been theorized only in the context of measurement practices (Reichenbach, 1927; Chang, 2001; Tal, 2012; Van fraassen 2008), can be generalized beyond the philosophy of measurement. Specifically, it will be shown that the problem is embodied in classificatory practices and that, consequently, historical coherentism is well suited to analyze these practices as well as metrological ones. As a case study, I look at a contemporary debate in phylogenetics, regarding the evolutionary origin of a newly identified archaeal phylum called *Methanonatronarchaeia*. Exploring this debate through the lens of historical coherentism provides a detailed understanding of the dynamics of the field and a foothold for critical analyses of the standard rationale used by practitioners.

1. Introduction

The objective of this paper is twofold. First, I present historical coherentism and argue that it is the best epistemological framework available to tackle the problem of coordination, an epistemic conundrum that arises with every attempt to provide empirical content to scientific theories, models or statements. Second, I argue that the problem of coordination, which has so far been theorized only in the context of measurement practices (Reichenbach, 1927; Chang, 2001; Tal, 2012; van Fraassen, 2008), can be generalized beyond the philosophy of measurement. Specifically, it will be shown that the problem is embodied in classificatory practices and that, consequently, historical coherentism is well suited to analyze these practices as well as metrological ones.

In short, coordination is the process by which theories, models or statements can be provided with empirical content. It involves linking these symbolic elements of knowledge-related endeavors with aspects of phenomena through the establishment of observational procedures. There is a *problem* with coordination in the sense that providing empirical content to theories always requires accepting a priori some assumptions that are constitutive of the theories we are trying to provide empirical content to. According to past approaches to this problem (e.g. conventionalism), it entails that empirical justification is circular. This diagnostic, however, is artefactual: it results from adopting an ahistorical perspective on empirical justification (van Fraassen, 2008).

Historical coherentism, in contrast, takes into account the sociohistorical dimension of scientific practice and justification, thereby turning the circle into a helix (see Fig. 1 in section 2.2). This sociohistorically inclined approach to epistemology holds that empirical justification is at the heart of scientific endeavors, but that empirical input is never sufficient to explain theory, model or statement acceptance or rejection. Historical coherentism does *not* solve the problem of coordination. It offers an epistemological framework to deal with the inevitable consequences of the said problem. It should be noted that I use the expression historical coherentism to refer to a heterogeneous body of work in philosophy of measurement that all share the core belief that epistemology must take into account sociohistorical factors when analyzing scientific endeavors, and that this follows from the problem of coordination.

The paper is divided in two parts, meant to address the two main objectives of the paper. First, I argue that historical coherentism is a useful framework to analyze any epistemic endeavors that must deal

* Corresponding author. Université Laval, Faculty of Philosophy, Québec, Canada.
E-mail address: francoisppl@gmail.com.

https://doi.org/10.1016/j.shpsa.2024.05.020
Received 26 May 2023; Received in revised form 13 March 2024; Accepted 23 May 2024
Available online 7 June 2024
0039-3681/

with the problem of coordination (i.e., any empirically driven field of knowledge). After having presented the problem in more details, section 2.1 explores the historical background that motivated the formulation of historical coherentism. In section 2.2, I describe the main conceptual tools I extracted from the work of Chang and van Fraassen. I will also suggest adding the notion of amalgam to the lot. In section 2.3, I clarify historical coherentism's relationship to objectivity and scientific progress.

In section 3.1, I generalize the problem of coordination, providing three new formulations (one to highlight its full reach and two to adapt it to classificatory practices). In section 3.2, I take the example of phylogenetics, a discipline part of evolutionary biology that aims at reconstructing the genealogical relationships between all living organisms, to illustrate how the problem of coordination is embodied in classificatory practices. In section 3.3, I analyze a recent debate in phylogenetics, regarding the evolutionary origin of a newly identified archaeal phylum. This analysis strengthens my argument according to which the problem of coordination can be generalized beyond measurement practices, but it also illustrates the explanatory power of historical coherentism, a toolbox for analyzing scientific practices while acknowledging the full complexity of empirical justification.

2. The problem of coordination and historical coherentism

Until now, the problem of coordination has been theorized as being related to metrological practices (Chang, 2001; Tal, 2012; van Fraassen, 2008). [1] In that context, it has been shown that any attempt to confirm a theory, model or statement through empirical observation (i.e. to coordinate it with empirical phenomena) requires assuming some parts of the theory. This whole conundrum has been synthesized by Chang as a four-steps process.

"(1) We want to measure quantity *X*.
(2) Quantity *X* is not directly observable by unaided human perception so we infer it from another quantity *Y*, which is directly observable.
(3) For this inference we need a law that expresses *X* as a function of *Y*, $X = f(Y)$.
(4) The form of this function f cannot be discovered or tested empirically, because that would involve knowing the values of both Y and X, and X is the unknown variable that we are trying to measure." (Chang, 2001, p.251).

It is a crucial situation for epistemology to address the problem of coordination since without coordination, any theoretical statement remains a piece of pure abstraction. Indeed, as van Fraassen puts it: "The theory would remain a piece of pure mathematics, and not an empirical theory at all, if its terms were not linked to measurement procedures." (van Fraassen, 2008, 115). Accounting for the ways in which scientific (read empirical) knowledge arises despite this apparent circularity should therefore be a priority for empirically minded approaches to epistemology and philosophy of science.

To illustrate the problem, consider the measurement of time. Time is a quantity that is not directly accessible to human perception (in any precise way). We can infer a length of time from the more direct observation of the periods of a clock (e.g., the Earth's rotation or the swing of a pendulum). The problem, in this situation, is that the establishment of the reliability of a clock involves assuming a specific relationship between time and the phenomenon that is to be used as a clock, namely that each period lasts the same amount of time. However, a clock, which is a physical phenomenon, cannot be shown to be temporally regular; to do so would require observing simultaneously two of its periods, a feat that is impossible. Hence, we must *assume* that the same periodic process always involves the same amount of time in order to have useable measurement procedures, which enable us to provide empirical content to various physical theories, models and statements. Reichenbach, nearly a hundred years ago, stressed the importance of this situation:

"Why is this determination impossible? Do not the laws of physics, for instance those of the motion of a pendulum, compel us to believe in the equality of the periods? It is true that the laws as described in textbooks suggest this belief; but if we ask ourselves where these laws come from, we shall find that they are obtained through observations of clocks calibrated according to the principle of the equality of their periods. The proof is therefore circular." (Reichenbach, 1927, p. 116).

The problem of coordination, in other words, is embodied when measurement procedures are being established and used to provide empirical content to theories, as these procedures necessitate a theoretical background that is part of what is to be empirically grounded. Before generalizing the problem of coordination beyond measurement practice (section 3.1), I will explore the body of literature that surrounds it in the philosophy of measurement. In section 2.1, I explore the main philosophical approaches of the 20th century that dealt with this conundrum. In section 2.2, I present historical coherentism.

2.1. *The philosophical background of historical coherentism: conventionalism and constructivism*

The framework I call historical coherentism was weaved to answer the weaknesses and echo the strengths of conventionalism (a form of foundationalism) as well as past versions of coherentism (e.g. constructivism, naively read). In this section, I review these approaches[2] and their relationship to the problem of coordination, to identify the core desiderata that guide historical coherentism. My analysis is based on work by Chang, van Fraassen and Tal.

Although in somewhat different terms, the problem of coordination was first theorized at the turn of the 20th century by authors such as Mach, Poincaré, Einstein and Reichenbach. Following the others's, Reichenbach's approach, called conventionalism, recognizes that the problem of coordination leads to conventions playing a significant part in the establishment of observational procedures. Indeed, the choice between two procedures (e.g. choice between mercury and air thermometers; Chang, 2004) is underdetermined by the theoretical dimension of the procedures and phenomena (which we can only access through theory-laden observational procedures). Given this underdetermination, there is always some leeway in the choice of a procedure, says Reichenbach, and acceptance of a given procedure always involves circularity, as expounded above.

This suggests that empirical knowledge rests on circular demonstrations of the validity of observational procedures, a circularity that is epistemically threatening according to some, including Reichenbach. When faced with this consequence of the problem of coordination, one way to react is to try and salvage the absoluteness of empirical knowledge by rejecting or boxing the menacing circularity. Reichenbach's conventionalism boxes circularity: while choice of procedures does entail circularity, that is not what empirical knowledge rests on. The relationship between knowledge and phenomena is mediated by *definitional principles* or, more colloquially, by the possibility for epistemic agents to establish definitions with allegedly straightforward relationships to phenomena. Because such definitions are semantic moves grounded in phenomena and because, according to Reichenbach, the

[1] Chang calls it the problem of nomic *measurement*.

[2] There is another important approach to the philosophy of measurement, namely, the Representational Theory of Measurement (Diez, 1997; Suppes et al., 1989). I do not discuss it here because it avoids exploring the practical aspects of measurement that make it possible to coordinate theory and phenomena (Mari, 2000, 2005). In other words, it neglects the problem of coordination.

semantic world lies outside the realm of knowledge, knowledge can be founded on something that is external to it: the semantic-phenomena relationship. In this light, it becomes clear that conventionalism is a central element of a broader foundationalist epistemology (Chang, 2007). According to that view, the absolute validity of empirical knowledge is salvaged despite the inevitability of epistemic complications resulting from coordination:

"A solution [to the problem of establishing the equal lengths of the periods of a clock] is obtained only when we [...] introduce the concept of a *coordinative definition* into the measure of time. The equality of successive time intervals is not a matter of *knowledge* but a matter of *definition*. As for spatial congruence, a certain rule must be laid down before the comparison of magnitudes is defined. This determination can again be made only by reference to a physical phenomenon; a physical process, such as the rotation of the earth, is taken as a measure of uniformity by *definition*. All definitions are equally admissible." (Reichenbach, 1927, pp. 116–117).

As numerous critiques have since then highlighted, this "solution" amounts to ignoring the problem instead of solving it or answering it in a meaningful way, if only because the establishment of semantic norms is just as underdetermined as the establishment of observational procedures leading to empirical knowledge (Quine, 1961).

More epistemologically significant to contemporary readers, however, is the fact that the conventionalist answer to the problem of coordination is completely ahistorical. As such, it translates a trend observed in foundationalist approaches more broadly. It underestimates the sociohistorical dimension of knowledge construction. Reichenbach suggests that a definitional core can be provided to empirical endeavors *ex nihilo*, through coordinative definitions, despite evoking (somewhat covertly; Reichenbach, 1927, as already cited above) the diachronic dimension of scientific practices. But this definitional core, crucial to Reichenbach's account, is also the result of past epistemic practices. Hence it fails to serve as an *absolute* foundation: the circularity is *not* adequately confined to the establishment of procedures by Reichenbach. As van Fraassen stresses, the theory-observation back-and-forth inherent to coordination has a deeper diachronic dimension neglected by Reichenbach:

"At the same time he means to be still actively inquiring into the empirical conditions under which such definitions can play the requisite role. What he did not do is change the ahistorical setting of his problem: the coordination is still apparently to be conceived of as possible in the absence of any previous such coordination. But is that possible at all? We have to ask more or less the same question again as before: how can such coordinative definitions be meaningfully introduced except in a historical context where there are some prior coordinations already in place? I submit that they cannot." (van Fraassen, 2008, 121).

Two ideas are inherent to van Fraassen's analysis of Reichenbach. First, even if a historical ground zero could be found, amounting to something like "the first coordination between a theoretical quantity and an aspect of a phenomenon", that ground zero would be so far back in time as to be irrelevant to our understanding of current epistemic practices. Second, such a moment might not be identifiable. It follows from this that ahistorical approaches to how theories gain their empirical content are misleading; empirical grounding is a phenomenon without absolute foundations in which the sociohistorical context always plays a crucial role.

This gives us a first philosophical breakthrough upon which historical coherentism was built: if sociohistorical factors influence scientific practice, then empirical inputs are only *a part* of all the factors that determine theory, model, or statement acceptance. The aim of epistemology is therefore to understand how all these factors give rise to an epistemic system with some degree of internal coherence, that involves the various dimensions of scientific practice and the context in which they take place. Again, van Fraassen's phrasing is eloquent:

"To be sure, this requirement of coherence is not simply one for logical consistency. Whether a sort of mechanism can be used to define the family of standard clocks depends on empirical regularities that may or may not obtain. The central coherence condition on the family of standard clocks, recall, was that if two are in coincidence, they run at the same rate, they run in synchrony. That is a matter of empirical fact. But if two sorts of mechanism satisfy this condition, there is no matter of fact as to which runs evenly, and a choice or convention alone can decide on one of them." (van Fraassen, 2008, 136).

This sets the stage for historical coherentism, which is rooted in the realization that empirical data is never sufficient to understand the scientific practices that surround a phenomenon. The phrase "is never sufficient" must be understood as implying that empirical data is often (but not always) part of the factors determining scientific work; it warrants empiricism but precludes naive empiricism.

Yet historical coherentism must avoid difficulties other than the sociohistorical blind spots of conventionalism, those highlighted by classical critiques of past forms of coherentism (Bender 1989; BonJour, 2017). For example, according to critiques of constructivism (who tend to make a strawman out of it), this approach denies the empirical world any influence on scientific practices, which could then be conceived as an entirely social practice (which it is) free of any contact to phenomena (which it is not). This view of constructivism is a misconception (Latour, 2003). As Chang aptly puts it: "Fixed points can be artificially created in the same way seedless watermelons can be created; these things cannot be made if nature will not allow them, but nonetheless they are our creations" (Chang, 2004, p. 49). This comparison highlights that when knowledge is built (e.g. the establishment of fixed points in thermometry), it is the result of complex interactions between the epistemic community and what Chang calls "nature" (which I would rather refer to as "phenomena").

Behind this misinformed critique of constructivism lies a serious worry for any form of coherentism. While it is important to understand the sociohistorical context that allows coordination and determines its modalities, it is also important, for scientists and philosophers studying scientific practice, to be able to understand scientific progress. An approach to epistemology that would be overly concerned with the internal dynamics of a scientific community could miss an important driver of scientific progress (however defined), namely, empirical support.

Tal illustrates this worry perfectly by discussing the difference of stability between two types of clocks. In both cases, modeling practices realized by a scientific community can dismiss the discrepancies between individual clocks of a given type. However, between corrections, one type of clock may remain much more stable than the other, such that its use can be considered as progress:

To illustrate this point, imagine that metrologists decided to keep the same algorithm they currently use for calculating UTC [Coordinated Universal Time], but implemented it on the human pulse as a standard clock instead of the atomic standard. As different humans have different pulse rates depending on the person and circumstances, the time difference between these organic standards would grow rapidly from the time of their latest correction. Institutionally imposed adjustments would only be able to bring universal time into agreement for a short while before discrepancies among different pulse-clocks exploded once more. The same algorithm that produces UTC would be able to minimize adjustments to a few hours per month at best, instead of a few nanoseconds when implemented with atomic standards (Tal, 2016, p. 317).

The sociohistorical dimension of science, while necessary to understand justificatory practices, is insufficient *if* it leads to neglecting how epistemic communities interact with phenomena. Consequently, epistemological analysis must take into account the complex interactions between and within research communities, their sociohistorical setting, *and their target phenomena* to understand how scientific practice changes across time.

This leaves us with two intertwined desiderata around which historical coherentism was weaved: first, epistemology and philosophy of sciences must take into account the sociohistorical dimension of knowledge construction; second, it must acknowledge that the contact between a community and phenomena plays an important role in scientific justification.

2.2. Historical coherentism: the framework

I use the expression historical coherentism[3] to refer to a heterogeneous literature developed in the philosophy of measurement. Although differing in many ways, the approaches of van Fraassen (1980, 2008), Chang (2004; 2007), Morrison (2009; see also Morrison & Morgan, 1999), Tal (2012; 2016), among others, hold that scientific justification is a sociohistorical process shaped by many factors. All these authors agree that the problem of coordination is an unescapable conundrum which explains the impossibility to rely on empirical grounding naively to understand the dynamics of science. The aim is not to solve the problem, but to take into account its consequences.

Synthesising what I consider to be the strongest contributions of all these authors, I suggest that historical coherentism is an approach to epistemology that aims at understanding how epistemic endeavors can lead to empirical knowledge, despite the important challenges of coordination. It also aims at understanding how and why these scientific practices change across time. In what follows, I explore what I take to be core conceptual tools of historical coherentism: Chang's notion of epistemic iteration, van Fraassen's synoptic vision, the focus on models advocated by various authors and my contribution, the notion of amalgam.

The heart of Chang's contribution to historical coherentism is the notion of epistemic iteration, around which other notions are knitted:

"Epistemic iteration is a process in which successive stages of knowledge, each building on the preceding one, are created in order to enhance the achievement of certain epistemic goals.…In each step, the later stage is based on the earlier stage, but cannot be deduced from it in any straightforward sense. Each link is based on the principle of respect and the imperative of progress, and the whole chain exhibits innovative progress within a continuous tradition." (Chang, 2004, p. 226; see also p. 45–48).

This provides a central model for representing scientific dynamics, the helix (see Fig. 1), where *stages of knowledge* are conceived as successive *iterations* of scientific practices. Whereas foundationalism suggests justification is circular, the helix model adds a temporal dimension. Instead of going in circles, we progress, but each successive iteration (each "circle") is similar to the previous one. The stability is illustrated by the link between every "circle" that makes it a helix instead of series of independent iterations. This model is meant to be heuristic. It oversimplifies the dynamics of scientific practice to emphasize that progress is achieved, but only by relying on past iterations. Understanding scientific practices thus amounts to understanding 1- what goes on in a specific iteration or stage of knowledge and 2- what changes from one iteration to the next as well as what remains stable and 3- what explains change and stability.

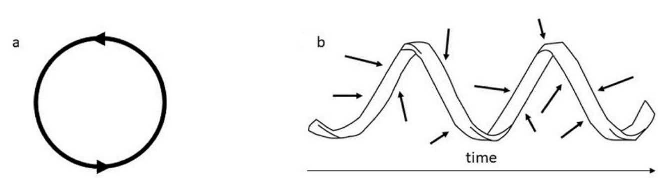

Fig. 1. This figure illustrates two distinct representations of scientific justificatory processes. In 1.a, scientific justification is conceived as a circular process. This is the kind of representation that Reichenbach, for example, offers as he abstracts away the temporal dimension of scientific practices. 1.b, illustrates the same back-and-forth process between theory and observations that is represented as a circle in 1.a, but the process now has a temporal dimension, which stretches the successive iterations into a helix. The arrows represent the variety of factors that influence scientific justification.

Stages of knowledge are not defined by Chang. He uses the term loosely to refer to any set of scientific practices that can be pragmatically identified to sustain informative analysis. Stages of knowledge can refer to general approaches to a given field of inquiry (e.g. Tree-thinking in evolutionary biology; O'Hara, 1997) or they can be more localized (e.g. a specific phylogenetic reconstruction realized by a group of researchers).

The dynamics underscored by the helix model result from, among other things, two vectors pressuring research in a sometimes-conflicting manner: the principle of respect and the imperative of progress. The principle of respect (Chang, 2004, pp. 43–44 and 256) is the idea that scientific practices must necessarily draw from previous iterations, from which they acquire, among other things, legitimacy. This principle makes sciences more or less conservative, depending on the weight of the principle of respect. The imperative of progress (Chang, 2004, pp. 44–46) is the engine for change, the injunction for scientists to do "better" than what has been done in the past. What is considered an improvement is highly context dependent. The perennial influence of both vectors implies that epistemic iterations follow one another, differ sometimes slightly and sometimes more drastically, but always rely on past iterations for structuring their justifications and for overcoming the challenges of coordination. In other words, the interaction between the two principles forms the backbone of the helix, of historical coherentism's outlook on scientific practice: "Each link [between iterations] is based on the principle of respect and the imperative of progress, and the whole chain exhibits innovative progress within a continuous tradition." (Chang, 2004, p. 46).

Once we accept the helix model as an abstract representation of the dynamics of scientific justification, the following questions can be answered to flesh accounts of specific sets of practices: what makes scientists hold some things to be improvements, while they respect certain aspects of past iterations so profoundly that they refuse to question them? Was new data mobilized and, if so, how was it generated? How were new ideas embedded in new iterations, how do they relate to older ideas and where do they come from? Etc. Answering these questions (and others) ultimately calls for more interdisciplinarity in scientific studies: historical, sociological, anthropological, and philosophical analyses will have to be brought together to understand the variety of factors that shape a specific helix.

Chang's approach can be complemented by van Fraassen's. Among the numerous and insightful contributions of van Fraassen in the context of the philosophy of measurement, one especially important for historical coherentism is the synoptic vision. Van Fraassen argues that epistemological approaches to scientific endeavors should fuse two perspectives on the object of inquiry: from within and from above (see van van Fraassen, 2008, p. 139). The first one refers to the importance of exploring the development of scientific practices from within its sociohistorical trajectory, as Chang does in his work. Looking at a series of iterations, we can thereby observe how the problem of coordination is being dealt with by practitioners, we can track gradual change and the factors that lead to these changes.

[3] I coined the expression, but I was strongly influenced by the one Chang uses to refer to his own approach, i.e. progressive coherentism. I use a different expression to stress that I refer to a broader literature.

Van Fraassen, however, insists on the need to cross this perspective with the one from above, which is the perspective that assumes the validity of a standardized theory's description of phenomena, avoiding any reference to the theory's own historical trajectory. In this ahistorical perspective, the validity of observational procedures is taken for granted, and so is the empirical support for the theory. This focuses our attention on the internal workings of a theory, on its logic, whose justificatory structure might very well be the target of epistemological analysis. This perspective, if taken alone, is problematic as it neglects the sociohistorical process at work which explains how and why a theory has changed and stabilized. Yet combining this perspective with the one from within the historical process provides insights into the justificatory structure of scientific endeavors, and their transformation across time.

For example, the synoptic vision can help us identify what I call *amalgams*. These are grouping of ideas that seem to be necessarily tied together to the point of being presented as a conceptual unit, even though they are historically and logically independent from one another. They can only be identified by joining two different perspectives together: the one from above, where the two ideas are fused together, and the one from within the historical process, where their potential independence is observed. If only the perspective from above is considered, there would be no reason to analyze the composite or even to describe it as such. If only the perspective from within is considered, the two ideas would simply be described as independent, thereby neglecting the intimacy of their contemporary interaction.

Such amalgams can be found everywhere and *are extremely useful to scientific practice*. For instance, the biological species concept is an amalgam. It refers to a community of descent (or a group of organisms in which there is or could be genetic flow), thereby implicitly tying a genealogical criterion to a taxonomic unit (Mayr, 1969). Historically and logically, however, it has been and could be otherwise (Zachos, 2016). Species can be defined based on ecological grounds, for example (Novick & Doolittle, 2021). The existence of amalgams implies that some of the conceptual associations that are required to infer interesting things about phenomena can be taken for granted without disrupting the whole scientific process. Indeed, if every association between ideas had to be dissected every time inferences were being realized, scientific endeavors would be much more tedious than what we observe in practice, to the point of inefficacy. Nonetheless, the identification of these amalgams is useful from an epistemological perspective, and that for at least two reasons: first, to describe scientific practices accurately; second, to identify potential levers for critical analysis.

Despite the fertility of the synoptic vision, van Fraassen's approach to the philosophy of measurement fits only partially what I want to make of historical coherentism. Specifically, his rationale is embedded in his interpretation of the semantic approach to scientific theories. According to it, models are parts of theories that are coordinated with phenomena. We have models on one side, phenomena on the other, and the two are to be coordinated through observational procedures. Hence, while van Fraassen stresses the theory-ladenness of observation, he still sees models as being distinct from the observation procedures themselves.

Coordinative procedures, however, involve models in a much more intimate manner than van Fraassen suggests. A measuring instrument, for example, is useless until the quantities it indicates are embedded in a model that provides symbolic meaning to the material dimension of the procedure. The process of coordination *is* the interpretation of physical phenomena through the lens of a model, a process that leads to measurement outcomes. Therefore, coordination cannot be used to establish the isomorphism of models and phenomena.

Consequently, it is best to keep away from the semantic interpretation of scientific theory. Following Tal (2016) and Mari (Mari & Giordani, 2012), I suggest banking on the pivotal role of models in coordination and weaving historical-coherentist analyses around them. Chang's epistemic iterations can be conceived as epistemic iterations of models, broadly construed.[4] Historical coherentism is meant to tackle how and why these models change across time (or why they do not), how different iterations are related to each other, how the use of a model is being justified, *and* how models are used to provide empirical content to broader theories, or specific statements. All of this is done by taking into account the sociohistorical dimension of scientific justification, a perspective that is necessary given the problem of coordination.

2.3. Historical coherentism, objectivity and scientific progress

In a 2019 paper, Isaac formulates a critique of Chang's, Tal's and van Fraassen's approaches to what he calls epistemic loops (the problem of coordination). His view is based on the notion of measurement success, which is used to uphold a form of realism regarding fixed points in metrology (such as the boiling point of water). Analyzing and rebutting his approach is beyond the scope of this paper, but one of his critiques of historical coherentism is important for the purpose of the present argument, because similar critiques are often aimed at various forms of coherentism.

That critique is centered on objectivity: "This coherentism in turn undermines *measurement realism*, the view that outcomes of successful measurement practices veridically represent objective (i.e., interest-independent) features of the world" (Isaac, 2019, p. 930; italics from the original). The use of the term "objective" is ambiguous. Indeed, the word is polysemous, with a family of meanings referring to methodological aspects of epistemic endeavors and others being ontologically driven (Lloyd, 1995). Here, the term "objective" is tied to "features of the world" and should therefore be understood ontologically. Yet the specification "interest-independent" is usually used to refer to a methodological imperative (that research be done free of the interests of researchers) or, in some cases, to the results of research upholding this methodological imperative. It can be argued that objective (in the sense of interest-independent) research leads to "veridical representation" of objective (e.g. existing independently of us) features of the world, but still the two notions should be clearly distinguished.

By conflating the two notions, Isaac's phrasing depicts historical coherentism as an approach that fails to account for objective knowledge altogether. This is misleading. It is true that the approach I defend in this paper dismisses the possibility of having knowledge about how things are, independently of us. According to me (some researchers tied to historical coherentism might disagree), the problem of coordination entails that we can only know about features of the world as they stand *in relation* to us. Having a God's eye view being impossible, scientific knowledge can only result from the complex processes of coordination between phenomena and theory. Rejecting ontological objectivity as an ideal is nonetheless coherent with the possibility of developing objective knowledge, where objectivity refers to methodological ideals. Historical coherentism and the gravitational pull of the problem of coordination

[4] Following Tal (2012), who himself follows Morrison (2009) and Cartwright (1999), I take models to be abstract representations of phenomena helpful for explanation and prediction regarding the concerned phenomena. Models are built upon various assumptions, which are acquired from overarching theories, from previous treatment of data or from (sometimes ad hoc) beliefs about phenomena. Models are produced by drawing on theories but are independent from them, to a certain extent (a model can be refuted without this having a major impact on theory acceptance, and vice versa).

require a notion of objectivity that rejects the "pure detachment" meaning of objectivity.[5]

Tal, for example, suggests that a statement communicates objective knowledge if the concerned community can be confident that it informs us about the object under inquiry:

"Prior to their representation by an idealized model, there is no way of testing whether different instruments measure the same quantity; any agreement or disagreement among their indications may be construed as coincidental and attributed to some local feature of the instruments or environments. It is only once their idiosyncrasies are idealized away in a mutually coherent fashion that instruments can be viewed as sources of objective knowledge about a common quantity, such as temperature or frequency." (Tal, 2017, p. 14).

Objectivity is obtained when the criteria of a community settle the validity of a knowledge statement. In this example, it means that they exclude the possibility that idiosyncrasies determine measurement outcomes. The strongest claim allowed by the problem of coordination is that objective knowledge is empirically supported knowledge that arises from standardized procedures, where standardization involves acceptance by the relevant community.[6] Objectivity is in this sense a methodological (or epistemological) notion but it is free from the ideal of detachment since it explicitly refers to a social process.

This may bring back worries regarding the importance of internal coherence being the main criterion for theory, model or statement acceptance. Indeed, what counts as objective knowledge might differ from one community to the next. Disagreements regarding what counts as objective knowledge do happen (even within individual communities, as the case study in section 3.3 highlights). The fact that a conception of objectivity acknowledges this is a strength. Moreover, acknowledging that objective knowledge differs from one community to the next still makes it possible to decide "objectively" between two alternative epistemic systems, although this requires adopting a specific viewpoint, or comparing many (hence the importance of van Fraassen's synoptic vision that invites us to always cross at least two distinct perspectives).[7] Air-based thermometer can be said to be objectively better than mercury-based ones, if one is willing to consider and accept the great variety of factors that led researchers to settle the debate thusly, or to adopt the perspective of subsequent iterations that sustained this conclusion (Chang, 2004).

Historical coherentism is all about studying why and how some practices and the associated knowledge are considered objectively better than others, thereby accounting for scientific progress. Its specificity, however, is to acknowledge and embrace the complexity of the process by which something ends up being considered objective knowledge or progress. This contrasts with accounts that would solve the debate by saying "this is objective knowledge because it reflects reality as it stands independently of us." Such approaches deny the existence of the problem of coordination; they short-circuit the complexity of epistemic endeavors and consequently, they misrepresent them.

3. Generalizing the problem of coordination and, in the process, historical coherentism

In section 2, I have presented the problem of coordination and historical coherentism. In what follows, I show that the problem can be generalized beyond measurement and is embodied in classificatory practices (section 3.1). A consequence of this is that historical coherentism is well suited to analyze classificatory practices.[8] In order to illustrate these claims, I will use phylogenetics as case study (sections 3.2 and 3.3).

3.1. *Generalizing the problem*

The problem of coordination can and should be reformulated more broadly, a generalization that highlights its import beyond measurement practices. I propose the following reformulation of Chang's four-steps description of the process (start of section 2).

(1) We want to observe phenomenon *X*.
(2) *X* is not directly observable by unaided human perception, so we infer it from another phenomenon *Y*, which is more directly observable.[9]
(3) For this inference we need to know the relationship between *X* and *Y*.
(4) The nature of this relationship cannot be discovered or tested empirically, because that requires the independent observation of *Y* and *X*, and *X* is the unobservable we are trying to observe.

Step 1 sets the objective (coordination through observation), steps 2 and 3 concern the establishment of observational procedures, while step 4 highlights problem. I argue that the situation is embodied by classificatory practices as well as metrological ones, a fact that has remained untheorized so far.

I define classificatory practices as epistemic endeavors that describe phenomena by dividing the world (into categories or classes of objects) and discriminating phenomena (in order to know which things belong to which categories), and then by reconstructing a system (or systems) of relationships relating the categories identified beforehand.[10] Accordingly, the problem of coordination, when embodied in classificatory practices, requires two distinct but intimately related formulations: one tied to the classification of objects *per se*, the other to the formation of systems. The two are intimately related because class-membership of specific objects can only be assigned in the context of an overarching classificatory system.

A first reformulation captures the categorization of objects.

(1) We want to observe relationship *X* between object *O* and a category *C*.

[5] Skepticism about this conception of objectivity is nothing new. Lloyd (1995), for example, has shown efficiently that even classical writings by authors such as Carnap, McDowell, Nagel or Searle questioned various aspects of the naive conception of objectivity that ties it to pure detachment from our cognition. The numerous challenges to objectivity as pure detachment, strengthened by the problem of coordination, provide grounds for doubting that it is relevant to criticize epistemological approaches (e.g. historical coherentism) based on the fact that they reject objectivity (conceived as pure detachment), as Isaac and others do.

[6] Tal's approach to epistemology echo many others, especially those tied to social epistemology, with Longino's social objectivity concept being an important benchmark (Longino, 1990). To manage values (Longino, 1990) and to standardize procedures of observation (Tal, 2017) are different but overlapping dimensions of the process leading to the creation of "objective" knowledge.

[7] This illustrates the complementarity of historical coherentism and some feminist approaches to philosophy of science, as paradigmatically heralded by Longino (1990).

[8] It should be noted that Chang has also tackled classificatory endeavors (e.g. Chang, 2016) using a methodology comparable to the one he uses for measurement. Similarly, Basso (2021) has drawn on a literature similar to the one I use to develop an approach to classification (in psychiatry). My approach differs from Chang's by introducing conceptual tools from other authors and from both his and Basso's by being centered on the generalization of the problem of coordination.

[9] "More directly observable" means, in this context, that we already have standardized observational procedures for it.

[10] This definition was greatly influenced by the work of Patrick Tort (1983; 1989), a historian and philosopher of biology whose extensive work on classification has yet to be translated to English.

(2) *X* is not directly observable by unaided human perception, so we infer it from another relationship *Y* between *O* and known or alleged members of *C*.
(3) For this inference we need to know the relationship between *X* and *Y*.
(4) The nature of this relationship cannot be discovered or tested empirically, because that requires the independent observation of *X* and *Y*, and *X* is the unobservable we are trying to observe.

In this first formulation, the relationship *X* will usually be that of belonging (we want to observe whether *O* belongs to *C*), but it could also be its opposite (not belonging to *C*). For example, we can ask whether an atom (*O*) belongs (*X*) to a given category (C), say an element. Because this is not directly accessible by human perception (even if we could see atoms directly, there would be no labels hovering over them), we must compare this atom with known or alleged members of the element to identify a similarity relationship (*Y*; based on the number of protons found in the nuclei, as suggested by the relevant theoretical framework). To confirm the validity of this inference, we would need to have independent access to class membership as well as to the atomic number of each atom, but class membership is what we are trying to provide a standardized observational procedure for.

A second reformulation captures the formation of systems.

(1) We want to observe relationship *X* between two classes of objects.
(2) *X* is not directly observable by unaided human perception, so we infer it from another relationship *Y*, which is more directly observable.
(3) For such inferences we need to know the relationship between *X* and *Y*.
(4) The nature of this relationship cannot be discovered or tested empirically, because that requires the independent observation of *X* and *Y*, and *X* is the unobservable we are trying to observe.

The case study, below, is focused on this second formulation, where *Y* is similarity and *X* is phylogeny. As I detail below, inferring phylogeny from similarity requires accepting a priori some aspects of the theory of evolution by natural selection, while enabling phylogenetic inquiries to provide empirical support for this same theory.

3.2. The problem of coordination in phylogenetics

The validity of the above generalization can be illustrated by showing how the problem of coordination is embodied in phylogenetics, a paradigmatic example of classificatory practices. To do so, we must first explore the relationship between phylogenetics and the overarching theoretical framework of evolutionary biology.[11] According to the standard view, natural selection drives evolution, meaning that random mutations are selected for (or against) and spread in (or disappear from) a population. As mutations accumulate across time, significant changes occur in the population. These changes are what evolution by natural selection (ENS) amounts to (Gayon, 1998; Godfrey-Smith, 2009; Huxley, 1942; Lewontin, 1970).

While it is fairly simple to establish that the theory of ENS informs us about changes occurring in specific populations (e.g. a population of fruit flies in a laboratory), the central claim of standard evolutionary biology is much stronger: the biological world as a whole has been shaped mostly by the slow but constant input of natural selection acting on biological populations, such that the theory of ENS can serve as an anchor for evolutionary explanations and as a background theory for all of biology (Gayon, 1998). Phylogenetics holds a privileged role in a successful consortium of scientific fields that, among other objectives, aims at providing empirical content to that stronger claim.

According to evolutionary biologists, the said theory makes predictions about the pattern of phylogenetic relationships that should be observed in the biological world. More precisely, they claim that if natural selection is indeed the main agent shaping evolution, then biological lineages should be related in a way that is best represented (i.e. modelled) in the form of a tree, i.e. the Tree of Life, the system of relationships that allegedly ties together all life forms.[12] Hence, if the specific hypothesis according to which the history of life is underscored by an arborescent phylogenetic system can be grounded empirically, this would offer the theory of ENS significant support.

According to traditional phylogeneticists, a good way to observe an arborescent phylogenetic system is to identify monophyletic groups, i.e. groups of organisms that include all and only the descendants of a common ancestor. Finding such groups in the wild is considered by many as an empirical contribution to the acceptance of Darwinism, broadly construed:

"Evolutionary theory predicts that monophyletic groups and only such groups emerge from various evolutionary processes termed speciation. They are composed of a common ancestral species and all of that species' descendants. […] Such groupings are sought because evolutionary theory predicts their existence. […] Thus, all truly monophyletic groups have the property of being composed of species, or higher taxa, who have exclusive, or unique, genealogical descent from a founder species. *Each higher taxon we hypothesize to be monophyletic stands as a singular confirmation of macroevolutionary theory because macroevolutionary theory predicts that such groupings should exist*." (Wiley and Lieberman 2011, 18; my emphasis).

The identification of monophyletic groups requires knowing the phylogenetic relationship between groups of organisms (e.g. species). This, however, is a serious challenge, as phylogenetic relationships are not directly accessible to human perception: they lie in the past. This situation embodies the problem of coordination.

(1) We want to observe phylogenetic relationships between taxonomic units (e.g. to identify monophyletic groups).
(2) Phylogenetic relationships are not directly observable by unaided human perception, so we infer them from similarity relationships, for which we have standardized observational procedures.
(3) For such inferences we need to know the relationship between phylogenetic and similarity relationships.
(4) The nature of this relationship cannot be discovered or tested empirically, because that requires the independent observation of phylogeny and similarity, and phylogeny is the unobservable we are trying to observe.

As a result, researches *must assume* the validity of the similarity-phylogeny relationship, which is embedded in the theory of evolution through the notion of heredity (Godfrey-Smith, 2009; Lewontin, 1970): like creates like, such that taxonomic units that are closely related phylogenetically will tend to be more similar than phylogenetically distant ones. A core principle of ENS is therefore to be accepted a priori if phylogenetics's grounding of the said theory is to be achieved.

This means that observations, in phylogenetics, are theory-laden, and it highlights the depth of the challenges met by scientists that

[11] Biologists distinguish phylogenetics, the establishment of phylogenetic relationships between genes and organisms, from phylogenetic systematics, i.e., the establishment of a nested hierarchy where groups form groups of higher ranks (species form genera that form families that form orders, etc.). Here, I focus on phylogenetics.

[12] The aim of this section is to illustrate how the problem of coordination is embodied in evolutionary biology. Accordingly, I take that which can be called traditional phylogenetics as a case study. Traditional phylogenetics is centered on tree-based models, but it must be noted that, since the turn of the 21st century, network-based models have been suggested as alternatives (Doolittle, 2000; Huson & David, 2006; Morrison, 2005).

hold empirical evidence dear. The contemporary standardized observational procedures for phylogenetic relationships is the result of a historical interplay between observations and theory, a series of iterations and refinements that involved, among other things, the phylogenetic interpretation of pre-Darwinian taxonomic systems (Inkpen & Doolittle, 2016; Panchen, 1992), the rejection of paraphyletic groups (Willmann, 2003), the development of classical genetics (Provine, 1989; Smocovitis, 1992), the quantification of similarity analysis (Hull, 2001; Sneath & Sokal, 1973) and the molecular turn (O'Malley, 2016). At none of these stages was empirical evidence for the theory of ENS *given*: "What now counts as simple passive measurement is a hard-won achievement" (van Fraassen, 2008, 125).

3.3. The placement of methanonatronarchaeia – a case study

Using very broad strokes, I illustrated, in the previous sections, how the problem of coordination is embodied in classificatory practices (through abstract formulations [section 3.1] and the example of the overarching rationale of phylogenetics [section 3.2]). In this section, I tackle a more specific case study in contemporary phylogenetics to showcase historical coherentism's conceptual tools.

I turn to a debate concerning the phylogenetic placement of the recently identified phylum of archaea called *Methanonatronarchaeia*, observed in environmental DNA (Sorokin et al., 2015). What is interesting, from an epistemological point of view, is that different groups of researchers (Aouad et al., 2019; Sorokin et al., 2017), starting with the same data, make very different hypotheses about its phylogenetic placement. In one case, the phylum is taken to be a sister-group of *Halobacteria* (Sorokin et al., 2017), while in the other, *Methanonatronarchaeia* is placed much deeper in the Tree of Life (Aouad et al., 2019). As will be detailed below, this difference is due to the methods used to analyze the data, i.e. to the procedures used to infer unobservable phylogenetic relationships based on the more readily observable genetic similarity. In the absence of a direct access to phylogenetic relationships, this situation illustrates the leeway in the establishment of procedures that is a consequence of the problem of coordination.

Sorokin et al. (2017, 2018, 2019) started with a phylogenetic (maximum-likelihood) analysis of 16S rRNA and concatenated alignments of ribosomal proteins of the *Methanonatronarchaeia* groups under study. The similarity between the two trees, further supported by "the fact that these trees conformed with the currently favoured solutions for difficult problems in archaeal phylogeny" (Sorokin et al., 2019, p. 560), led the authors to state that *Methanonatronarchaeia* form a sister-group of *Halobacteria*. *Methanonatronarchaeia* would accordingly be considered an intermediate evolutionary step between archaeal methanogens and *Halobacteria*.

In contrast, Aouad et al. (2019), using the same supermatrix of ribosomal proteins, established a much deeper placement for the archaeal lineage. This difference is explained by the removal of fast-evolving sites, which, as they argue on the basis of past research, tend to generate phylogenetic noise (Aouad et al., 2018). The alternative placement they suggested was also tested within two larger supermatrices, one composed of more markers (Aouad et al., 2018) and the other of a larger sampling of methanogenetic entities (Borrel et al., 2019). Their inclusion of more gene markers challenges a traditional focus of the discipline, which privileges 16S rRNA. All of this led Aouad et al. to suggest that the hypothesis of a sister-group including *Halobacteria* and *Methanonatronarchaeia* might be an artefact of tree-reconstruction methods.

Sorokin et al. (2019) responded with three arguments. First, they claim that supressing fast evolving site may indeed remove phylogenetic noise, but it can also lead to the loss of informative sites. They thus challenge a central assumption of standard phylogenetics according to which more stable sites are always carriers of better phylogenetic information (interestingly, the assumption they challenged is traditionally used, *mutatis mutandis*, to focus on markers such as 16S rRNA). Second, they highlighted that the larger matrixes are composed of proteins that are more susceptible to horizontal gene transfer (more on this below). Third, they remind their readers that larger matrixes cannot account for or dispel the results that were attained through phylogenetic analysis of 16S rRNA. They concluded by calling for more sampling of *Methanonatronarchaeia* with the hope that it might resolve the debate.

There are good reasons to doubt that more data can solve such disputes, strengthening the claim central to historical coherentism according to which empirical data is never sufficient to settle theory, model or statement acceptance or rejection. The assumption that species phylogenies can be positively reconstructed from genetic information, once we gather enough data, is challenged by the epistemic situation we here observe. Use of different gene markers leads to different phylogenetic reconstructions, and this should be expected given that different genes have different phylogenetic histories, some different from that of the species we observe them in. So, while we could speculate about the possibility of there being a fact of the matter regarding the phylogenetic placement of groups of organisms like *Methanonatronarchaeia*, we are forced to admit that we have no direct access to it and that as long as we rely on similarity to infer this placement, there will always be various possibilities available to us. The best we can hope for is to establish statistical tendencies in the forest of trees generated by gene-based phylogenetic reconstructions (Koonin et al., 2021), but even this fails to deny the fact that focus on different genes or the use of different procedures for similarity analysis will yield different results.

In the present case study, this situation is illustrated by the fact that both groups of researchers challenge some of the assumptions that are constitutive of standard phylogenetic procedures to defend the validity of their respective hypotheses. The debate, in other words, is not about these hypotheses and their validity. It is about the underlying methods used to infer phylogenetic relationships based on genetic similarity.

The justificatory structure at work can be represented using the helix model. The two approaches are competing stages of knowledges. In both cases, justification relies on past iterations of the field. For example, Sorokin et al. explicitly rely on past phylogenetic results to argue that their procedure is the right one (as cited above, they consider coherence with past phylogenetic reconstructions to be an argument supporting their hypothesis). Aouad et al. refer to past research to justify their removal of fast-evolving sites from the supermatrices. *More importantly, both groups of researchers retain the bulk of the inferential process typical of standard phylogenetic practices*: the methods used for similarity analysis are standard, and the phylogenetic inferences realized by both groups are tree-based (I will get back to this shortly).

This basic and unavoidable fact follows from the problem of coordination (absolute ahistorical justification is impossible; historically grounded justification is the rule), but it also means that justification is not circular. It is best described by a helix in which the principle of respect translates the continuity between successive iterations. Concurrently, the imperative for progress echoes the fact that both groups of researchers challenge standard practices in some ways or another (Sorokin et al. suggest that the use of less stable sites may be informative; Aouad et al. include a wide range of gene markers, beyond traditional ones such as 16S rRNA). Such challenges can lead to significant changes in scientific practices, across time.

Further critical analysis of this case study reveals a situation in which respect towards past beliefs and past iterations might be misleading. Both groups of researchers use tree-based models to infer phylogenetic relationships (Sorokin et al. used maximum-likelihood analysis to reconstruct branching phylogenies, while Aouad et al. mobilized Bayesian in addition to maximum-likelihood analysis). In other words, they assume, as most practitioners in the field do (Delsuc et al., 2005; Felsenstein, 2004; O'Hara, 1997; Parks et al., 2018; Wiley & Lieberman, 2011), that branching phylogenies are to be expected and found in the biological world. This belief justifies the use of tree-based models.

In the case of *Methanonatronarchaeia*, this assumption is surprising, given that the analyses of both sets of researchers provide good reasons to challenge the standard tree-based model for phylogenetic reconstruction. First, two distinct hypotheses for the placement of *Methanonatronarchaeia* both have empirical support, which means that a single tree might not be able to illustrate the whole evolutionary history of the phylum (both hypotheses might translate relevant dimensions of the history of the lineage, e.g., the history of distinct genes). Second, there might be convergence involved, as demonstrated by Aouad et al. Convergence is incompatible with branching phylogenies yet might be worth representing. Third, some of the genes used for inferential purposes are susceptible to lateral gene transfer. Representing these transfers requires including reticulation in phylogenetic graphs, which, again, challenges the validity of tree-based models. Hence, there are good reasons to use alternative models, yet the researchers involved in the debate failed to challenge the validity of tree-based models. The situation can be explained by drawing on the notion of amalgam.

As a reminder, amalgams are groupings of ideas considered to be necessarily intertwined, despite being historically and logically independent from one another. Both groups of researchers working on *Methanonatronarchaeia* uphold the validity of an amalgam characteristic of traditional phylogenetics, i.e. the intimate association of ENS and divergence. The unwillingness to consider alternatives to the tree may very well be explained by the conflation of divergence and ENS.

In Darwin's articulation of the theory of ENS, the principle of divergence plays a central role. This principle states that individuals that differ more from the mean have better chances of success (Darwin did not use statistical language, but the idea remains the same). These diverging individuals, because of their characteristic differences, tend to compete less directly with the bulk of the population; selective pressures are thereby less stringent for them. Given the ubiquity of natural selection and the resulting importance of the principle of divergence, says Darwin, one should expect the lineages that thrive to diverge ever more from one another (thanks to heredity, they accumulate ever more diverging traits). This claim plays a key role in Darwin's argument: natural selection leads to diverging evolution. *Therefore*, it can explain the taxonomic systems that were available at the time (which were then reinterpreted as phylogenetic systems that match the tree-based model and the principle of divergence).

The amalgam according to which ENS entails diverging patterns, implicit in the debate regarding *Methanonatronarchaeia*, reaches far beyond the said debate. It is explicit in various textbooks. For example, Felsenstein claimed that "[p]hylogenies, or evolutionary trees, are the basic structures necessary to think clearly about differences between species." (Felsenstein, 2004, xix). Wiley and Lieberman also vehiculate this amalgam: "Methods that explicitly test hypotheses of the descent of species have resulted in rigorously tested phylogenetic trees" (Wiley and Lieberman 2011, 1). O'Hara, in a programmatic paper meant to synthesize the main tenets of standard phylogenetic practices, claims that: "[i]f we seek to understand common causes acting in evolution then the replicates we need to examine are not species, but the evolutionary events that are of interest in a particular study, *and this can only be done by plotting those events on a tree*" (O'Hara, 1997, p. 325; my emphasis). Hence, it is not surprising that some research published in leading scientific journals (some of the papers by Sorokin et al. and Aouad et al. mentioned above were published in *Nature*) assumes the validity of the composite of ideas according to which ENS and tree-based models go together, to the point of amalgaming the two.

Adopting the synoptic vision makes it possible to contrast the use of this amalgam by contemporary researchers, and its history. Divergence was a mean for Darwin to coordinate his theory, centered on natural selection, with phenomena. It enabled the reinterpretation of similarity relationships between organisms and species, upon which available taxonomic system were built, as phylogenetic relationships, in coherence with postulated consequences of ENS. By looking at the work of many contemporary evolutionary biologists, we see that the cause (ENS) and the postulated consequence (divergence), are conflated. Hence, challenging divergence and tree-based models or simply considering alternative models has been and is still viewed by many as a direct challenge not only to these models, but also to the theory of ENS (Forterre, 2012; Koonin & Wolf, 2009; Merhej & Raoult, 2012; Raoult, 2010; Raoult & Koonin, 2012).

This takes us back to the principle of respect. The presence of an amalgam and its a priori acceptance by researchers is not itself a problem. The problem of coordination entails the need for some degree of respect for past iterations, and this respect can be actualized by the acceptance of specific sets of ideas and relationships between them. Empirical justification requires this. Nonetheless, respect can be misplaced: it might harm scientific practice when it is granted to underserving associations of ideas. The amalgam linking tree-based models and ENS is a case in point, as recent research in evolutionary biology provides good reasons for rejecting it: ENS and divergence do *not* go hand-in-hand (Bapteste & Boucher, 2009; Doolittle, 2000; Huson & David, 2006; Huson et al., 2010; Morrison, 2005, 2010; Papale et al., 2020). Future debates, regarding *Methanonatronarchaeia* (or other objects of inquiry) would therefore benefit from an in-depth exploration of alternatives to the tree-based representations of evolutionary history. Critical analysis can foster progress.

Finally, going back to the start of our epistemological journey, this example shows that there is no way for scientists to "solve" or overcome once and for all the problem of coordination: justification is never absolute. Scientists continuously manage coordination, re-adapting their procedures to new data, but also to novel stance towards data sets, new methods or algorithms, new tools, etc. Far from being limitative or paralyzing, the conundrum resulting from coordination forces researchers to innovate constantly: acknowledging it is an opportunity and an important condition for progress to occur.

4. Conclusion

The aim of this paper was to present and defend historical coherentism as a sound epistemological approach to tackle not only metrological practices, but also classificatory ones. In section 2, I presented the problem of coordination. This allowed me to highlight the epistemological importance of this phenomenon (the problem of coordination) for understanding the processes of empirical justification. It also allowed me to motivate the use of historical coherentism as an epistemological framework: the problem of coordination forces us to acknowledge the sociohistorical dimension of scientific justification and to develop a sophisticated account of objectivity and scientific progress. Following Tal (2017), I argued that objectivity characterizes knowledge that arises from standardized scientific practices.

Stepping away from metrological discussions, I showed, in section 3.1, that the problem of coordination can be generalized and reformulated in the context of classificatory practices. In section 3.2, I highlighted its embodiment in the general epistemic structure of phylogenetics. In 3.3, I refined the example by looking at a specific case study, namely, the debate regarding the phylogenetic placement of *Methanonatronarchaeia*. This illustrated the explanatory power of historical coherentism and its conceptual tools. Applying them to this specific case of classificatory practices shows that the scientific dynamics at work in contemporary phylogenetics do fit the helix model and that, accordingly, the scientific process can be analyzed in terms of the principle of respect and the imperative of progress. By identifying an amalgam (linking ENS and divergence) in the epistemic structure characterizing a significant part of contemporary phylogenetics, I have suggested that misplaced respect may impair scientific practices. In a near future, historical coherentism should be used to provide rigorous comparative analyses of tree-based and network-based phylogenetics, the two main alternatives available in the field. Ideally, this should be done by mobilizing interdisciplinary resources, going beyond philosophical analysis. Future work will also need to test the generality of the problem of coordination beyond phylogenetics.

CRediT authorship contribution statement

François Papale: Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I wish to thank Eran Tal, Gordon McOuat, Tyler Brunet, Stefan Lindquist, and Ford Doolittle for insightful comments on the paper. I am also grateful for the very sustained feedback provided by Cassandre Ville, Pierre-Olivier Méthot, Molly Kao and François Duchesneau. Special thanks are due to Frédéric Bouchard and Éric Bapteste for their support throughout the process from which this manuscript emerged. Part of the work was done with the financial support of The Gordon and Betty Moore Foundation (GBMF9729, https://doi.org/10.37807/GBMF9729), of the New Frontiers in Research Fund (grant NFRFE-2019-00703), and of the Social Sciences and Humanities Research Council of Canada (756-2023-0275).

References

Aouad, M., Borrel, G., Brochier-Armanet, C., & Gribaldo, S. (2019). Evolutionary placement of Methanonatronarchaeia. *Nature Microbiology, 4*(4), 558–559. https://doi.org/10.1038/s41564-019-0359-z

Aouad, M., Taib, N., Oudart, A., Lecocq, M., Gouy, M., & Brochier-Armanet, C. (2018). Extreme halophilic archaea derive from two distinct methanogen class II lineages. *Molecular Phylogenetics and Evolution, 127*(October), 46–54. https://doi.org/10.1016/j.ympev.2018.04.011

Bapteste, E., & Boucher, Y. (2009). Epistemological impacts of horizontal gene transfer on classification in microbiology. *Horizontal gene transfer: Genomes in flux*.

Basso, A. (2021). From measurement to classificatory practice: Improving psychiatric classification independently of the opposition between symptom-based and causal approaches. *European Journal for Philosophy of Science, 11*(4), 104. https://doi.org/10.1007/s13194-021-00424-y

The current state of the coherence theory: Critical essays on the epistemic theories of keith lehrer and laurence BonJour, with replies. In Bender, J. W. (Ed.), *Philosophical studies series* (Vol. 44), (1989). Dordrecht ; Boston: Kluwer Academic Publishers.

BonJour, L. (2017). The dialectic of foundationalism and coherentism. In J. Greco, & E. Sosa (Eds.), *The blackwell guide to epistemology* (1st ed., pp. 117–142). Wiley. https://doi.org/10.1002/9781405164863.ch4.

Borrel, G., Adam, P. S., McKay, L. J., Chen, L.-X., Natalia Sierra-García, I., Sieber, C. M. K., Letourneur, Q., et al. (2019). Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nature Microbiology, 4*(4), 603–613. https://doi.org/10.1038/s41564-019-0363-3

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge, UK. New York, NY: Cambridge University Press.

Chang, H. (2001). Spirit, air, and quicksilver: The search for the" real" scale of temperature. *Historical Studies in the Physical and Biological Sciences, 31*(2), 249–284.

Chang, H. (2004). Inventing temperature: Measurement and scientific progress. In *Oxford studies in philosophy of science*. Oxford ; New York: Oxford University Press.

Chang, H. (2007). Scientific progress: Beyond foundationalism and coherentism. *Royal Institute of Philosophy Supplement, 61*, 1–20.

Chang, H. (2016). The rising of chemical natural kinds through epistemic iteration. In *Natural kinds and classification in scientific practice* (pp. 53–66). Routledge.

Delsuc, F., Brinkmann, H., & Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics, 6*(5), 361–375. https://doi.org/10.1038/nrg1603

Diez, J. A. (1997). A hundred years of numbers. An historical introduction to measurement theory 1887-1990-Part II: Suppes and the mature theory. Representation and uniqueness. *Studies In History and Philosophy of Science Part A, 28* (2), 237–265.

Doolittle, W. F. (2000). Uprooting the tree of life. *Scientific American, 282*(2), 90–95.

Felsenstein, J. (2004). *Inferring phylogenies*. Sunderland, Massachusetts: Sinauer Associates.

Forterre, P. (2012). Darwin's goldmine is still open: Variation and selection run the world. *Frontiers in Cellular and Infection Microbiology, 2*(August). https://doi.org/10.3389/fcimb.2012.00106

Gayon, J. (1998). *Darwinism's struggle for survival: Heredity and the hypothesis of natural selection*. Cambridge: Cambridge University Press.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. New York: Oxford University Press.

Hull, D. L. (2001). The role of theories in biological systematics. *Studies in History and Philosophy of Biol & Biomed Sci, 32*(2), 221–238. https://doi.org/10.1016/S1369-8486(01)00006-1

Huson, D. H., & David, B. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution, 23*(2), 254–267. https://doi.org/10.1093/molbev/msj030

Huson, D., Rupp, R., & Scornavacca, C. (2010). *Phylogenetic networks: Concepts, algorithms and applications*. Cambridge: Cambridge Univ. Press.

Huxley, J. (1942). *Evolution: The modern synthesis*. London: Allen & Unwin.

Inkpen, S. A., & Doolittle, W. F. (2016). Molecular phylogenetics and the perennial problem of homology. *Journal of Molecular Evolution, 83*(5–6), 184–192. https://doi.org/10.1007/s00239-016-9766-4

Isaac, A. M. C. (2019). Epistemic loops and measurement realism. *Philosophy of Science, 86*(5), 930–941. https://doi.org/10.1086/705476

Koonin, E. V., Makarova, K. S., & Wolf, Y. I. (2021). Evolution of microbial genomics: Conceptual shifts over a quarter century. *Trends in Microbiology, 29*(7), 582–592. https://doi.org/10.1016/j.tim.2021.01.005

Koonin, E. V., & Wolf, Y. I. (2009). Is evolution darwinian or/and lamarckian? *Biology Direct, 4*(1), 42.

Latour, B. (2003). The promises of constructivism. In *Chasing technology: Matrix of materiality, ihde, don, 27–46. Indiana series for the philosophy of science*. Indiana: Indiana University Press.

Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics, 1*, 1–18.

Lloyd, E. A. (1995). Objectivity and the double standard for feminist epistemologies. *Synthese, 104*(3), 351–381.

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton, N.J: Princeton University Press.

Mari, L. (2000). Beyond the representational viewpoint: A new formalization of measurement. *Measurement, 27*(2), 71–84. https://doi.org/10.1016/S0263-2241(99)00055-X

Mari, L. (2005). The problem of foundations of measurement. *Measurement, 38*(4), 259–266. https://doi.org/10.1016/j.measurement.2005.09.006

Mari, L., & Giordani, A. (2012). Quantity and quantity value. *Metrologia, 49*(6), 756–764. https://doi.org/10.1088/0026-1394/49/6/756

Mayr, E. (1969). The biological meaning of species. *Biological Journal of the Linnean Society, 1*(3), 311–320.

Merhej, V., & Raoult, D. (2012). Rhizome of life, catastrophes, sequence exchanges, gene creations, and giant viruses: How microbial genomics challenges Darwin. *Frontiers in Cellular and Infection Microbiology, 2*, 113.

Morrison, D. A. (2005). Networks in phylogenetic analysis: New tools for population biology. *International Journal for Parasitology, 35*(5), 567–582. https://doi.org/10.1016/j.ijpara.2005.02.007

Morrison, M. (2009). Models, measurement and computer simulation: The changing face of experimentation. *Philosophical Studies, 143*(1), 33–57. https://doi.org/10.1007/s11098-008-9317-y

Morrison, D. A. (2010). Phylogenetic networks in systematic biology (and elsewhere). In *Research advances in systematic biology*. Mohan, R., 1–48. India: Trivandrum.

Morrison, M., & Morgan, M. S. (1999). Models as mediating instruments. In M. S. Morgan, & M. Morrison (Eds.), *Models as mediators* (pp. 10–37). Cambridge University Press. https://doi.org/10.1017/CBO9780511660108.003.

Novick, A., & Doolittle, W. Ford (2021). Species' without species. *Studies In History and Philosophy of Science Part A, 87*(June), 72–80. https://doi.org/10.1016/j.shpsa.2021.03.006

O'Hara, R. J. (1997). Population thinking and tree thinking in systematics. *Zoologica Scripta, 26*(4), 323–329.

O'Malley, M. A. (2016). Histories of molecules: Reconciling the past. *Studies In History and Philosophy of Science Part A, 55*(February), 69–83. https://doi.org/10.1016/j.shpsa.2015.09.002

Panchen, A. L. (1992). *Classification, evolution, and the Nature of biology. Cambridge [england]*. New York, NY, USA: Cambridge University Press.

Papale, F., Saget, J., & Bapteste, É. (2020). Networks consolidate the core concepts of evolution by natural selection. *Trends in Microbiology, 28*(4), 254–265. https://doi.org/10.1016/j.tim.2019.11.006

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*. https://doi.org/10.1038/nbt.4229. August.

Provine, W. B. (1989). Progress in evolution and meaning in life. In M. Nitecki (Ed.), *Evolutionary progress* (pp. 49–74). Chicago: University of Chicago Press.

Quine, W. Van O. (1961). Two dogmas of empiricism. In *From A logical point of view* (2nd ed., pp. 20–46). Cambridge, Massachussets: Harvard University Press.

Raoult, D. (2010). The post-darwinist rhizome of life. *The Lancet, 375*(9709), 104–105. https://doi.org/10.1016/S0140-6736(09)61958-9

Raoult, D., & Koonin, E. V. (2012). Microbial genomics challenge Darwin. *Frontiers in Cellular and Infection Microbiology, 2*. https://doi.org/10.3389/fcimb.2012.00127

Reichenbach, H. (1927). *The philosophy of space and time*. Courier Dover Publications.

Smocovitis, V. B. (1992). Unifying biology: The evolutionary synthesis and evolutionary biology. *Journal of the History of Biology, 25*(1), 1–65. https://doi.org/10.1007/BF01947504

Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical taxonomy: The principles and practice of numerical classification*. San Francisco: W.H. Freeman.

Sorokin, D. Y., Abbas, B., Mitchell, G., Pimenov, N. V., Sukhacheva, M. V., & van Loosdrecht, M. C. M. (2015). Methanogenesis at extremely haloalkaline conditions in the soda lakes of kulunda steppe (altai, Russia). *FEMS Microbiology Ecology, 91*(4). https://doi.org/10.1093/femsec/fiv016

Sorokin, D. Y., Makarova, K. S., Abbas, B., Ferrer, M., Golyshin, P. N., Galinski, E. A., Ciorda, S., et al. (2019). Reply to 'evolutionary placement of Methanonatronarchaeia. *Nature Microbiology, 4*(4), 560–561. https://doi.org/10.1038/s41564-019-0358-0

Sorokin, D. Y., Makarova, K. S., Abbas, B., Ferrer, M., Golyshin, P. N., Galinski, E. A., Ciordia, S., et al. (2017). Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nature Microbiology, 2*(8). https://doi.org/10.1038/nmicrobiol.2017.81

Sorokin, D. Y., Merkel, A. Y., Abbas, B., Makarova, K. S., Rijpstra, W. I. C., Koenen, M., Sinninghe Damsté, J. S., Galinski, E. A., Koonin, E. V., & van Loosdrecht, M. C. M. (2018). Methanonatronarchaeum thermophilum gen. Nov., sp. Nov. and 'candidatus methanohalarchaeum thermophilum', extremely halo(Natrono)Philic methyl-reducing methanogens from hypersaline lakes comprising a new euryarchaeal class Methanonatronarchaeia classis nov. *International Journal of Systematic and Evolutionary Microbiology, 68*(7), 2199–2208. https://doi.org/10.1099/ijsem.0.002810

Suppes, P., Krantz, D. H., Luce, R. D., & Tversky, A. (1989). *Foudations of measurement* (Vol. 2). San Diego and London: Academic Press, 3 vols.

Tal, E. (2012). *The epistemology of measurement: A model-based account*. University of Toronto. PhD Thesis, Ph. D. dissertation.

Tal, E. (2016). Making time: A study in the epistemology of measurement. *The British Journal for the Philosophy of Science, 67*(1), 297–335.

Tal, E. (2017). A model-based epistemology of measurement. N. Mössner and A. Nordmann. In *Reasoning in measurement* (pp. 233–253). New York: Routledge/Taylor and Francis Group.

Tort, P. (1983). *La Pensée Hiérarchique et l'évolution. Les Complexes Discursifs*. Paris: Aubier Montaigne.

Tort, P. (1989). *La raison classificatoire: Quinze études*. Editions Aubier.

van Fraassen, B. C. (1980). The scientific image. *Clarendon library of logic and philosophy*. Oxford : New York: Clarendon Press ; Oxford University Press.

van Fraassen, B. C. (2008). *Scientific representation: Paradoxes of perspective*. Oxford University Press.

Wiley, E. O., & Lieberman, B. S. (2011). *Phylogenetics: Theory and practice of phylogenetic systematics*. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118017883

Willmann, R. (2003). From haeckel to hennig: The early development of phylogenetics in German-speaking europe. *Cladistics, 19*(6), 449–479. https://doi.org/10.1111/j.1096-0031.2003.tb00383.x

Zachos, F. E. (2016). *Species concepts in biology*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-44966-1