DRAFT VERSION – PLEASE DO NOT CITE

Theory Construction as Meta-Induction and Meta-Induction as a Form of Induction

Guy Hetzroni

Abstract
Theoretical methods that provide the basis for proposed physical theories are often developed and justified based on their similarity to methods that were applied in the construction of established theories. The paper argues that naturalist epistemology constrains the notions of similarity that should be considered as 'projectable' in such meta-inductive arguments, in a way that can contribute both to understanding past theories and to forming rational judgments on the pursuitworthiness of proposed projects. The account is demonstrated using two examples: differential equations and invariance arguments.

1. Introduction

Can the empirical success of a theory provide justification for other theories that are constructed using similar methods? Is it possible to draw justifiable and general methodological lessons from successful theorizing? These questions are particularly important in scientific fields in which experimental testing lags behind theory construction. Their relevance is highlighted in the context of current fundamental physics, where the pursuitworthiness of various theoretical approaches is debated before empirical evidence is at hand. This situation presents various challenges, both to science and the philosophy of science. Current discourse is characterized by diverging opinions on the general compatibility of the practice of parts of theoretical physics (e.g., string theory, multiverse theories) with traditional standards of empirical science (Dawid, 2013; Ellis and Silk, 2014; Hossenfelder, 2018; Dardashti et al., 2019), as well as on the content and justification of principles and theoretical virtues (e.g. Crowther and Rickles, 2014; Holman, 2014; Borrelli and Castellani, 2019; Crowther, 2021; Fischer, 2023). These debates bear on the issue of pursuitworthiness, at the intersection of epistemology of science and science policy.[1]

The current paper aims to contribute to the discussion by sketching a general philosophical framework that would allow for a more refined discussion in the compatibility of theoretical methods and accepted standards of empirical science. It aims to show how these standards constrain the conditions under which the past success of theoretical methods should be projected to future theorizing. This issue is framed in terms of the compatibility of meta-inductive reasoning with naturalism.

[1] See Shaw (2022) and other articles in the same issue.

Meta-induction is a way of understanding the applicability of various forms of theoretical reasoning as an extension or extrapolation of theoretical considerations from accepted theories to suggested ones. Dawid (2013, 2016) presents meta-inductive arguments as a means of confirmation and assessment: new empirical evidence for certain theories also improves the epistemic status of other theories that were constructed using similar methods. This picture suggests that such evidence can also urge scientists to further develop and apply similar methods, making meta-induction not only a form of assessment, but also a methodological approach to theory construction. In this broad sense, the appeal to physical principles in theory construction, and also to theoretical virtues and mathematical-conceptual frameworks, can all be regarded as forms of meta-induction, as far as the appeal is based on the applicability of these principles, virtues and concepts to theories whose empirical success has been established.

This is a very broad conception of meta-induction. Fruitful application of it requires some criteria that would tell us why some meta-inductive arguments are better justified than others. (The same is true in the case of Dawid's Bayesian account, see Smolin, 2014; Chall, 2018; Dawid, 2022.) This paper aims to present such criteria by addressing the question of which meta-inductive arguments can be considered rational and justifiable from a naturalist point of view, i.e. one that aims for evidence-based justification and dismissal of *aprioristic* reasoning. This view is briefly presented and formulated as a working hypothesis in §2. The core of the paper is the discussion in the question of which mete-inductive inferences can be considered 'projectable', which is presented in §4, with analogy to more familiar naturalist accounts of projectability of standard inductive inferences, that are presented in §3.

Clearly, justifying meta-induction is relevant not only to well-debated projects in cutting-edge theoretical physics. To put meta-induction on secure grounds, it is instructive to look at better-understood case studies. Why do scientists expect that differential equations similar to ones used in classical physics would be applicable to other domains? Does the fact that invariance arguments contributed to the formulation and unification of certain laws of interaction justify the other ongoing attempts for unification based on invariance, and if so, are some attempts more justifiable than others? §5 uses these examples to demonstrate the applicability of the presented account of the projectability of meta-inductive arguments.

2. Epistemic naturalism

Often contrasted with epistemic foundationalism, anthropocentrism or certain forms of idealism, naturalism is identified with a cluster of ideas aiming for continuity of philosophy and science. Here we are concerned with the compatibility of certain scientific methods and naturalist epistemology.[2] The form of naturalism that we need here would serve to ensure that the view of science that is implicit in scientific practice is by itself a view that adheres to scientific standards. Meta-induction in particular is a reflective practice, which has to rely, either implicitly or explicitly on a certain view of science. It is carried out from a third person's perspective on the scientific process, explaining its empirical success and taking epistemological lessons from it. Roughly speaking,

[2]This project is different from the project of naturalizing metaphysics, which aims to read ontology from our theories (e.g., Ladyman et al., 2007).

this practice is compatible with naturalism if scientific process is understood on equal footing with scientific understanding of natural processes.

A form of naturalism that meets this desideratum is that of 'subject naturalism', comprehensively presented and advocated by Price (2011), who defines it as 'the philosophical viewpoint that begins with the realization that we humans (our thought and talk included) are surely part of the natural world' (p. 5). In particular, scientific theories and theoretical concepts go under 'our talk and thoughts'. Thus, if the best way of gaining knowledge and understanding of the natural world is an inquiry that is guided and constrained by scientific standards, subject naturalism would imply that these standards are also the key to understanding scientific process itself, and account for its successes and failures.

It is not the aim of this paper to engage in detail with the question of what are these scientific standards. It is sufficient for our purposes to take a broad view according to which scientific understanding of natural processes would privilege empirical evidence and, furthermore, favour *a posteriori* justifications and explanations over *a priori* ones and appeal to specific contingencies over general necessities. The naturalist hypothesis we shall adopt would imply that these priorities should also shape our understanding of scientific progress. It is, apparently, not trivial to reconcile these standards with the kind of mathematical reasoning that characterizes meta-inductive inferences in contemporary physics. The aim of this paper is to show that this naturalist hypothesis therefore constrains which meta-inductive inferences should be considered acceptable.

3. Projectability judgments

Faced with the question of what makes certain meta-inductive inferences more justifiable then others, the discourse on scientific induction presents us with a useful analogy, most clearly expressed in Goodman's new riddle of induction. Goodman (1983) notes that past observations can be characterized in terms of infinitely many different predicates, that can accordingly give rise to many contradicting inductive inferences. In a famous example Goodman points out that a set of observations compatible with the hypothesis that all emeralds are green is equally compatible with the hypothesis that all emeralds are grue (where the predicate *grue* 'applies to all things examined before [a certain time in the future] t just in case they are green but to other things just in case they are blue', p. 74). Accounting for the inductive inference 'all emeralds are green' thus requires a criterion for the choice of this particular predicate 'green' from among many possibilities. Some attempts to resolve or dissolve the problem apply constraints over some aspect (e.g. the time-dependence) of the predicates (e.g., Jackson, 1975). On a broader reading, however, Goodman's riddle reveals a deeper issue about scientific induction, showing that inductive inferences require criteria that distinguish between projectable and non-projectable classifications that are equally compatible with the data and give rise to inductive arguments whose logical form is identical. Making a choice from among the possible classifications is an essential part of any inductive argument, yet the this choice is never enforced by the evidence. Quine (1969) presents a perspective, centered around the notion of 'natural kinds', according to which these classifications arise gradually from more basic notions of similarity, that are expressed, refined and become entrenched in the process that gives rise to accepted and established theories. The grouping and characterization of useful classifications is described by Quine, following Goodman's conception of entrenchment, as a 'second-order induction'. Quine, however, stresses that this process is conducted with the aim to reflect a real and objective relation of similarity in the world. Every inductive inference therefore depends on a projectability judgment, according to which a certain theoretically-based classification imposed on the data is preferable over others that are equally compatible with it.

Boyd (1991) presents a more detailed account of projectability judgments in the same spirit. One central aim of his account is to facilitate an argument for realism, an issue beyond the scope of the current paper. The more relevant part, however, is Boyd's appeal to epistemic naturalism in his analysis of projectability judgments. The appeal to natural kinds serves the purpose of establishing projectability judgments on an *a posteriori* basis. These judgments are based on *background theories*, i.e., on both the theoretical content and the observational content of accepted theories that provide the background knowledge for the inductive inference. The belief that classifications underlying inductive inference reflect natural kinds, i.e., that they originate in contingencies found in nature, therefore allows scientific induction to be understood as rational by science's own standards.

4. The Projectability of Meta-Inductive Arguments

We are now in a position to reformulate our original question in terms of projectability of meta-inductive arguments. Is there a way to identify notions of similarity between theoretical methods that give rise to projectable classifications of them, and thus to projectable meta-inductive arguments? This is not an easy task. There are clearly many ways to understand, reconstruct and classify theoretical methods. The extrapolation of theoretical considerations to the generation of new theories can be done in numerous ways, and there is not always any clear notion of similarity. To approach the question we turn to examine whether the naturalist treatment of the projectability of inductive arguments can be extended to cases of meta-induction.

There are, *prima facie*, good reasons to doubt the possibility of such an extension. Theoretical methods can be based on an intricate and often inseparable combination of empirical, theoretical, mathematical, and hypothetical considerations. We would not expect such considerations to fall under 'natural kinds'. Furthermore, the question of projectability is often cast in terms of confirmation: under what conditions does a body of data support a particular inductive step? It would make less sense to ask under what condition does a body of data support the use of a certain theoretical method. The question of the projectability of meta-induction therefore departs in some ways from the discussion on induction. Yet, there are also similarities. Like in the case of induction, the question here concerns judgments of theoretical plausibility and the way they reflect on the rationality of the scientific process. Given the past success of a set of theoretical arguments that fall under a certain classification, when will it be rational to expect that other arguments falling under this classification would lead to empirically successful theory? Our aim here would be to show that the naturalist discussion of induction is also relevant to this question.

This task requires a more detailed account of what meta-inductive arguments actually are and and their workings in the scientific process. To accommodate them in the naturalist view, we shall define meta-inductive inferences as an inductive-like inferences, for

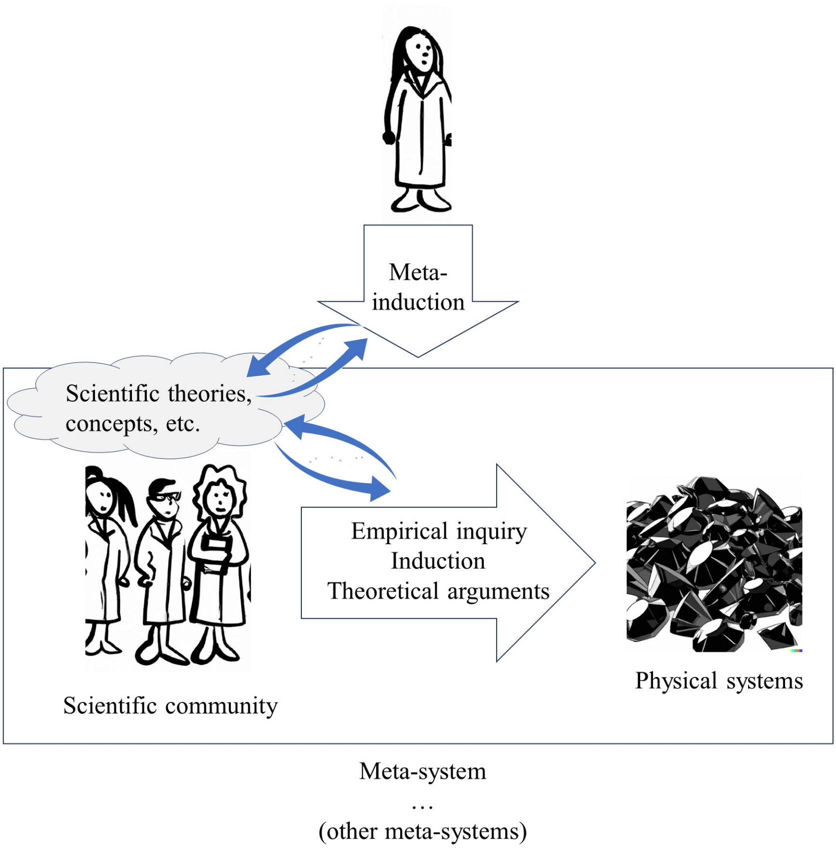

Figure 1. A meta-inductive inference .

which the relevant target system includes not only the physical systems that are the subject of inquiry (e.g., emeralds, or elementary particles), but also the scientific community that studies those physical systems. In other words, the process we are concerned with is theory construction, and its components include both the relevant physical objects and the community studying them (including the community's knowledge and the theories being employed). It is assumed that the community is a scientific community in the sense that theories are primarily judged based on their empirical adequacy.

On this view, the agent making the meta-inductive inference describes the scientific process from a third person perspective (Figure 1). For the sake of clarity, let us distinguish between the physical system of interest, and the *meta-system* that includes the system and the scientific community.[3]

Let us now assume that in the course of the scientific process a theoretical argument has been employed, giving rise to a new theory. Later, the predictions of the theory are put to experimental test. After this process has been carried out several times (with some

[3]The physicalist would further argue that the meta-system is also a physical system. This stronger statement is not an essential part of the current analysis.

theories turning out to be empirically adequate, and others inadequate) a meta-induction is in place.[4] In this description, standard induction and theoretical considerations are employed from within the meta-system (i.e., by the scientific community, in the context of the relevant background knowledge) while the meta-inductive inference is made from an external point of view. A meta-inductive inference is based on identifying a similarity among theoretical considerations that have lead to empirically adequate theories, and will argue in favour of theories that were not yet put to experimental test but were constructed using similar theoretical considerations. The argument can go further, and articulate a way of applying similar considerations in theory construction, thus suggesting new theories.

According to the naturalist view of scientific induction, projectability judgments are *a posteriori*, theory-dependent classifications of scientific terms, identified (in a cautious, revisable manner) with natural kinds. In the case of meta-induction, the relevant classifications are of theoretical arguments. These arguments are part of the overall meta-system, that includes the scientific community and the physical system being investigated. There would be many possible classifications, depending of various epistemic, formal and physical considerations. In particular, some classifications would be based on the 'human' side of the meta-system, i.e. on the knowledge, beliefs and concepts employed by the scientific community. Other classifications can be based on conjectured properties of the physical systems. Notably, in many arguments there may not be a clear-cut separation between these types. Due to theoretical holism, a given argument is often understood on one reading as based on certain concepts that constitute the theoretical framework, and on another reading as relying on physical properties of the described systems. A projectability judgment would have to take a stand here, classifying a set of arguments, in a certain way based on a particular reading.

Our concern is with the justification of judgments on the projectability of meta-inductive arguments. The name of the game is empirical adequacy. Projectability judgments of meta-inductive arguments aim to characterize arguments that have lead to empirically adequate theories in the past, under a classification that can be projected to new theoretical arguments aimed at producing new empirically adequate theories. Thus, a projectability judgment is a belief on what it is exactly that was 'done right' in previous cases of successful theorizing, an attempt to draw a (cautious and tentative) lesson from past success and failures. With this aim in mind, it is clear that a judgment on the likelihood of a certain theoretical method lead to an empirically adequate theory should not be based simply on a conceived similarity of the suggested application of the method to previous successful cases. In order for the judgment to be rational from a naturalist point of view, the similarity should be between contingent factors that may have contributed to the past success of the methods. Explicating what are these factors and how they may have contributed is part of a good projectability judgment.

Moreover, in the presented scheme of the scientific inquiry, it is clear that the empirical adequacy of a theory is not just a property of the theory, nor is it a property of the observations. Empirical adequacy is clearly a relational property, associated with a

[4]While the applicability of this simplified account to different cases can be questioned on several grounds, it does capture a significant aspect of the cases of interest, which is the belated arrival of empirical evidence after a theory was constructed.

relation between the theory and the system. In other words, empirical adequacy cannot be confined to either the conceptual (or human) side of the meta-system, nor to the physical side, it characterizes a relation between the concepts employed by the community and the physical world, and it must be accounted for as such. Any classification of theoretical methods based only on the kinds of systems they describe is therefore non-projectable, nor is a classification that is based only on concepts and ideas. In order to be projectable, a classification of theoretical methods has to account for the way in which particular methods lead to an empirically successful theory, and to do so it has to refer to the way theoretical concepts relate, either to evidence, or to some aspects of the studied physical system (observable or hypothesizes ones).

In §2 we motivated a naturalist view of the process of scientific inquiry, that aims to explain the outcomes of the inquiry in scientific standards that privilege explanantia based on evidence, contingencies, and *a posteriori* knowledge. Based on this view, we arrive now at the following necessary projectability condition for meta-induction: *In order to be projectable, a classification of theoretical arguments has to characterize them based on the way theoretical concepts relate to available evidence or to other form of scientific knowledge about contingent and specific features of the different situations.*

One major implication is that classifications of theoretical arguments that are solely based on the formal concepts they apply are not projectable.[5] As we shall see in the examples below, this necessary condition is far from being trivial. In fact it can filter both philosophical accounts and suggested scientific accounts of the success of certain theoretical methods and have potential implications on theory construction and debated philosophical issues.

Finally, a short comment on the retrospective nature of projectability judgments is in place. A projectability judgment is an attempt to draw a lesson from past successes and failures. This would require some appeal to the historical process of theorizing. Since our concern is with the rationality of science, a projectability judgment would better not be completely detached from actual considerations employed by scientists. Yet, the process of historical development is usually a messy combination of various elements, including irrelevant historical coincidences, time and place dependent sociological factors, and various contingencies related to the scientific inquiry itself (e.g., what measurements were technically possible at a given time, or which experiments have actually been made). A projectability judgment is thus also a judgment on which factors are the ones that are relevant and scientifically valuable. It would have to reconstruct a clean, idealized (and obviously partial) picture of the actual development in which the theoretical arguments have been employed. Just like in the case of scientific induction, projectability judgments here have to be approached with a fallibilist attitude, they can be revised as new evidence is accumulated and new knowledge is gained. Accordingly, our reflective understanding of past theories and the reasons for their success and failures is likely to change over time. In this way, naturalism implies that our

[5]This is in agreement with Steiner's claim that 'for the naturalist, anthropocentric hypotheses are "unprojectible". [...] That is, we cannot (if naturalists) argue that we are simply "doing the same thing" when the criterion for "same" is an anthropocentric one.' (1998, p. 143). Steiner, who did not refer explicitly to meta-induction, regarded any form of mathematical reasoning as essentially anthropocentric. According to the view presented here, the existence of this non-projectable classification of arguments, does noft preclude the existence of a projectable characterization of the same arguments.

philosophical understanding of past science depends on current theories, and can change alongside with them.

5. Two examples

Differential equations and invariance arguments are major theoretical methods whose applicability was projected to various domains based on prominent cases of successful applicability. While providing a thorough account for each case is a task for another day, this section will briefly use these examples to demonstrate the implications of the presented naturalist conception of the projectability of meta-inductive arguments.

5.1. Differential equations

The motion of bodies and the changing current in electric circuits are both successfully described in physics in theories based on ordinary differential equations that depend on the time variable (t-ODEs, for brevity). During the 20th century this mode of theory construction was transported to other fields such as population dynamics, atmospheric science and economics. To understand these examples in terms of meta-induction, we can think of the physical theories using t-ODEs as background theories for the construction of theories in other fields. A projectability judgment in this context would answer the following question: what exactly in the successful application of particular physical theories makes us expect that similar methods should be applied to certain other fields? Any possible answer to this question would depend on a way of understanding the workings of differential equations in physical theories.

One possible answer would be focused on the notion of differential equations itself. We simply learn from physics that differential equations are a useful description, and we should use them more. Referring only to the conceptual (or human) side of the meta-system, it is this kind of answer that does not satisfy the necessary condition of §4. What this paper comes to stress, is that this kind of answer is not rational from a naturalist point of view, i.e., it seems to account for the past success of science in a way that is in tension with scientific standards. In this sense, the above understanding of differential equations is on a par with manifestly non-naturalist accounts, such as McCloskey's (1991) analysis of the applicability of differential equations. McCloskey views science as a special case of rhetoric, not inherently different from other forms of rhetoric.[6] Dynamical models and differential equations are derived, in McCloskey's account, from metaphors and are being used to produce stories, where both metaphors and stories are understood as ways of arguing. This view (see McCloskey, 1983) is not in conflict with the formal nature of the theories, but rather with the view that they are empirically based. From the point of view of projectability, this account is not very useful. It does not seem to provide any basis for a distinction between the processes for which differential equations are likely to be useful, such as changes in the fish population in the Upper Adriatic (Volterra, 1928), and many other cases in which it will be useless (such as the changes in number of eggs John Doe got in his fridge). From

[6]McCloskey aims to present an alternative to positivism and modernism, primarily in the context of economics, but, *inter alia* criticizes the more general naturalist view that science is characterized by unique methods and by having prediction as a goal (McCloskey, 1983).

the naturalist perspective, this is directly related to the disengagement from empirical guiding standards.

It is instructive to contrast this perspective with a very different one by Einstein (1927). In reflecting on Newton's legacy he wrote: 'The differential law is the only form which completely satisfies the modern physicist's demand for causality' (p. 255). Differential equations express the way in which 'the state of motion of a system gives rise to that which immediately follows it in time' (p. 255). Whether this view manifests a naturalist conception of the projectability of the applicability of differential equations depends on what is the origin of the demand for causality. The naturalist reading would find it in background theories, i.e., regard causality as a contingent theory-dependent standard. Such standards are generally conceived as the basis of projectability judgments.[7] More precisely, there would be two elements to this view. First, we believe due to our background knowledge on each of the different cases that there are interdependent rates of change of certain quantities (both in the background physical theories and in the other theories being constructed using ordinary differential equations). Second, these quantities have measurable impact in which the scientific community is interested, even though they may provide, at most, a very partial description of the systems. For example, the analysis by Volterra (1928) applies differential equations to describe the population size of two species because they are conjectured to be interdependent due to predator-prey interactions (and also large enough to be approximated by continuous variables), and also because these are the particular quantities on which the coupling of the physical system to the human community depends (the population sizes of a relevant fish species have measurable impact relevant for fishing).

Thus, the projectability of t-ODEs, i.e. the conditions for their capacity to provide helpful description, does not strictly depend on the described systems, nor on the conceptual framework employed by the scientific community. Some relevant facts are relational, such as the aspects of the system that are of interest to the community, the available knowledge about it and the measurement capabilities. Appeal to these relations between the community and the system is necessary in any projectable account of the applicability of differential equations.

5.2. Invariance arguments

Invariance arguments have a central place within the toolkit of the contemporary theoretical physicist. The invariance arguments that were applied in the formulation of contemporary theories of gravity and particle physics can be described in various ways. Some locate the core of invariance arguments in the aim to unify by enlarging the symmetry group. Some identify it with the reduction of interactions to bundle geometry. Other approaches are based on a collection of more specific principles, e.g. the localization of global symmetries or the construction of theories whose symmetries guarantee (via Noether's theorem) local conservation of certain quantities. According to those possible characterizations, different enterprises apply different symmetry principles. While

[7] A different reading is to regard the causality requirement as a relativized *a priori*, e.g. in the spirit of Cassirer (1956). A major difference between this view and many naturalist views is the Kantian commitment to a hierarchical structure of knowledge. The compatibility of such structure with the naturalist hypothesis presented in §2 depends on the justification that is given for the hierarchical structure and seems to deserve further examination.

in some cases there is an overlap between these considerations, in other cases they lead to different, contradicting theories.[8] This situation makes the issue of the projectability of invariance arguments a pertinent question within theoretical physics, which is particularly pressing given the large number of suggested theoretical modifications for accepted theories.

From the perspective presented in Section 4, the localization of a gauge group, fibre-bundle structure and unification by enlarging the symmetry group, are concepts that cannot support by themselves a projectable meta-induction as they do not relate explicitly to to evidence or knowledge about specific features of the situation. Reference to conserved quantities and Noether currents seems potentially more projectable. A different characterization is given in Hetzroni and Read (2024), according to which the apparently formal localization of a global symmetry is in fact a straightforward way of taking into account the locality of the evidence that supports the interaction-free theory, by adding fields that locally determine relevant structure (such as inertial frames or preferred representations). This characterization of invariance arguments is based on the interface between theoretical concepts and evidence, and can therefore potentially explain the way the arguments lead to coupling terms in an empirically successful theories. The characterization is thus projectable. More specifically, it shows that the similarity between the role of general covariance in general relativity and different applications of the gauge argument is not a similarity between the different interactions, nor it is a similarity between the mathematical structures that represent them. It is a similarity between the patterns of inference through which suggested interaction terms where constructed based on local evidence.

6. Conclusion

This paper reflects on questions debated in current theoretical physics, but it bears relevance to other fields in which theory is pulling ahead of experiment and theories are often empirically tested only long time after they have been developed. In such situations meta-induction is a significant form of justification and method of theory construction. This paper suggests to view meta-induction as an inductive inference that is about scientific inquiry itself. I argued that in such situations, the capacity of a theoretical method to lead to an empirically adequate theory should be regarded as an explanandum, and that judgments on the projectability of meta-inductive arguments should be based on notions of similarity that apply to the explanantia. In this account constant reflection on past theories contributes to the assessment of suggested theories and to constraining the possibilities in mathematically-guided theory construction. It should be interesting to explore the implications of this naturalist conception of meta-induction on debated cases such as string theory and multiverse theories, and also on the broader issue of pursuitworthiness.

[8] One example concerns the attempts to apply the concept of gauge to gravitational theories. Utiyama (1956) suggested to construct general relativity by gauging the Lorentz group. Shortly after, Kibble (1961) noted that the the requirement that *every* global symmetry should be made local leads to the gauging of the Poincaré group, implying a (practically undetectable) deviation from general relativity in the coupling of spinors. Similar unification programs further extend the gauge group (Hehl et al., 1995). Understanding these theories geometrically can lead to coupling prescriptions that would further alter the empirical content (Delhom, 2020).

References

Borrelli, A. and Castellani, E. (2019). The practice of naturalness: A historical-philosophical perspective. *Foundations of Physics*. *49*(9), 860–878.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*. *61*(1/2), 127–148.

Cassirer, E. (1956). *Determinism and Indeterminism in Modern Physics: Historical and Systematic Studies of the Problem of Causality*. New Haven: Yale University Press.

Chall, C. (2018). Doubts for Dawid's non-empirical theory assessment. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. *63*, 128–135.

Crowther, K. (2021). Defining a crisis: the roles of principles in the search for a theory of quantum gravity. *Synthese*. *198*(14), 3489–3516.

Crowther, K. and Rickles, D. (2014). Introduction: Principles of quantum gravity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. *46*, 135–141.

Dardashti, R., Dawid, R. and Thébault, K., editors. (2019). *Why Trust a Theory? Epistemology of Fundamental Physics*. Cambridge University Press.

Dawid, R. (2013). *String theory and the scientific method*. Cambridge University Press.

Dawid, R. (2016). pp. 191–205 in *Modelling Non-empirical Confirmation*. edited by Ippoliti, E., Sterpetti, F. and Nickles, T. Springer International Publishing.

Dawid, R. (2022). Meta-empirical confirmation: Addressing three points of criticism. *Studies in History and Philosophy of Science*. *93*, 66–71.

Delhom, A. (2020). Minimal coupling in presence of non-metricity and torsion. *European Physical Journal C*. *80*(8), 1–17.

Einstein, A. (1927). The mechanics of Newton and their influence on the development of theoretical physics. *Die Naturwissenschaften*. Reprinted in Einstein (1954) pp.253–261.

Einstein, A. (1954). *Ideas and opinions*. Crown Publisher, New York.

Ellis, G. and Silk, J. (2014). Scientific method: Defend the integrity of physics. *Nature News*. *516*(7531), 321.

Fischer, E. (2023). Naturalness and the forward-looking justification of scientific principles. *Philosophy of Science*. p. 1–10.

Goodman, N. (1983). *Fact, Fiction, and Forecast*. Harvard University Press. fourth edition.

Hehl, F. W., McCrea, J., Mielke, E. W. and Ne'eman, Y. (1995). Metric-affine gauge theory of gravity: field equations, Noether identities, world spinors, and breaking of dilation invariance. *Physics Reports*. *258*(1), 1 – 171.

Hetzroni, G. and Read, J. (2024). How to teach general relativity? *The British Journal for the Philosophy of Science*.

Holman, M. (2014). Foundations of quantum gravity: The role of principles grounded in empirical reality. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. *46*, 142–153.

Hossenfelder, S. (2018). *Lost in math: How beauty leads physics astray*. Hachette Book Group.

Jackson, F. (1975). Grue. *The Journal of Philosophy*. *72*(5), 113–131.

Kibble, T. W. (1961). Lorentz invariance and the gravitational field. *Journal of Mathematical Physics*. *2*(2), 212–221.

Ladyman, J., Ross, D., with John Collier and Spurrett, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford University Press.

McCloskey, D. N. (1983). The rhetoric of economics. *Journal of Economic Literature*. *21*(2), 481–517.

McCloskey, D. N. (1991). History, differential equations, and the problem of narration. *History and Theory*. *30*(1), 21–36.

Price, H. (2011). *Naturalism without mirrors*. Oxford University Press.

Quine, W. v. O. (1969). Natural kinds In *Essays in Honor of Carl G. Hempel: A Tribute on the Occasion of his Sixty-Fifth Birthday*, Rescher, N. (eds). Springer Netherlands. Dordrecht. pp. 5–23.

Shaw, J. (2022). On the very idea of pursuitworthiness. *Studies in History and Philosophy of Science*. *91*, 103–112.

Smolin, L. (2014). String Theory and the Scientific Method. *American Journal of Physics*. *82*(82), 1105––1107.

Steiner, M. (1998). *The applicability of mathematics as a philosophical problem*. Harvard University Press.

Utiyama, R. (1956). Invariant theoretical interpretation of interaction. *Physical Review*. *101*(5), 1597.

Volterra, V. (1928). Variations and Fluctuations of the Number of Individuals in Animal Species living together. *ICES Journal of Marine Science*. *3*(1), 3–51.