Testing Theories as a Concern of the Sociology of Science

Iulian GRIGORIU[1]

Abstract
In the context of recent transformations in science, marked by Big Data, computational simulations, and artificial intelligence, the concept of theory testing requires re-evaluation. From the universalist ideal of Popperian falsifiability, we are shifting toward forms of testability that are local, situational, and institutionally mediated. The central issue lies in clarifying what it means to test a theory in contemporary sciences, a problem that generates tension between universal and local forms of testability. The aim of this article is to provide arguments in favour of a philosophy of "situated testability," grounded in an interdisciplinary methodological approach. Following a critical analysis of how testing is currently understood and practiced in science, I argue for a mode of testability that is adequate to the complexity and diversity of contemporary epistemologies, and I seek to answer the question "To what extent does scientific testing become an institutional and social practice?" Contrary to the universalist and deductivist ideal advanced by traditional philosophy of science, I contend that testability today emerges as a situated, infrastructural, and contextual practice. I thus advocate a pluralist and critical orientation, grounded in contextual experimentalism and in the localization of scientific laws.

Keywords: *Situated testability; contextual robustness; methodological pluralism; data-centric science; distributed testability; institutional validation;*

1. The Contextual Robustness of Scientific Theory Testing

Classical concepts of testing derive from Popperian epistemology, where a scientific theory was required to be falsifiable, that is, to generate empirically testable predictions that could potentially be refuted. In the age of simulations and Big Data, this criterion is being eroded. Models increasingly provide probabilistic scenarios rather than clear predictions. Algorithms detect patterns without our necessarily knowing what exactly is being tested. As a result, testability becomes operational and local, depending on the structure of data, access to computational infrastructure, and the institutional objectives of research.

Today, the epistemic conditions of testing have changed radically, in contexts where theories are no longer formulated as deductive hypotheses but are instead understood as adaptive systems, algorithms, or simulations. Moreover, replicability, another traditional pillar of validation, has been

[1] PhD, Senior Lecturer, "Dunărea de Jos" University of Galaţi, Romania, e-mail: iulian.grigoriu@ugal.ro

undermined due to technical and institutional difficulties, such as lack of access to data, dependence on software infrastructures, and variability in algorithmic implementation.

In this context, a new criterion emerges, contextual robustness. A theory or model is considered valid not merely through the possibility of strictly replicating a phenomenon, but through its ability to produce consistent results under different conditions. In this sense, testing becomes a social and institutional practice, what counts as testable is determined by criteria imposed by funding agencies, academic journals, infrastructures, and disciplinary norms. Thus, scientific validation today shifts from being solely an epistemological issue to one that is also political and organizational.

2. Contextual Experimentalism and the Localization of Scientific Laws

Beginning in the twentieth century, the philosophy of science gradually moved away from the positivist and universalist model, in which laws were conceived as necessary descriptions of nature, valid everywhere and at all times. This shift was supported by authors such as Nancy Cartwright, Sandra Mitchell, and Sabina Leonelli, who observed that scientific laws operate under specific, controlled, and localized conditions.

Nancy Cartwright argues that the laws of physics are abstract instruments that function in idealized contexts (for example, in the laboratory). In the real world, phenomena are far too complex, so laws are often approximate or even false outside those restricted conditions. She introduces the notion of "causal capacities" (1983), suggesting that a theory is testable insofar as these capacities can be isolated and controlled. Consequently, the laws of physics should be understood as approximations rather than faithful descriptions of reality. Cartwright formulates the distinction between the "phenomenological" and the "theoretical" in two distinct registers (1983). For philosophers, the opposition follows the line between observable and unobservable; for physicists, it separates the description of phenomena from theoretical foundations. Within this framework, Cartwright assumes a specific form of anti-realism. She accepts phenomenological laws, since they describe what happens, but rejects the truth-claim of fundamental laws, whose explanatory power is achieved at the cost of descriptive adequacy. The onto-gnoseological stake is explicit, theoretical laws in physics do not assert truth, even if they spectacularly organize and unify disparate domains.

Realism, by contrast, maintains that if a theory explains phenomena well, it is probably true, or at least approximately true. Bas van Fraassen challenges this line of reasoning by asking why should the explanatory power of a theory guarantee that it is true. He points out that the history of science is full of theories that provided excellent explanations yet proved to be false, such as phlogiston theory, the ether theory, or Newtonian mechanics in certain contexts. To believe that a theory is true merely because it explains well is an unjustified extrapolation. This constitutes a critique of inference to the best explanation (IBE), a principle frequently invoked by scientific realists.

Here van Fraassen develops what he calls "constructive empiricism", an anti-realist philosophical position which holds that a scientific theory must "save the phenomena", that it must be empirically adequate, while the question of its truth is irrelevant. "To accept a theory is not to believe it is true, but only to believe it is empirically adequate." (van Fraassen, 1980, p. 12)

Cartwright takes up and expands van Fraassen's question, why should explanatory power guarantee the truth of a hypothesis? The only case in which the explanation-truth relation gains strength is in causal explanation, when we describe the concrete process producing a phenomenon. In such cases, accepting the explanation involves ontological commitment to the causes invoked, but not to the truth of the fundamental equations instrumentally used in the process. The classic example of Crookes's radiometer and Maxwell's reinterpretation shows how rational belief in theoretical entities (gas molecules, tangential stresses) can be supported by experimentally determined causal roles without conferring truth upon the abstract equations of kinetic theory. According to Cartwright, explanatory success depends exclusively on lower-level causal principles and on concrete phenomenological laws (1983, p. 10).

The structural tension between causal and theoretical explanation thus becomes, for Cartwright, evidence of the falsehood of fundamental laws in their applications. On the one hand, in order to compose causes from different domains, physicists need formulas that indicate the contribution of each cause; on the other hand, those very formulas succeed only by ignoring the interferences of other laws operating in real situations. From this, she concludes that "the laws of physics do not state the facts" when causes are combined, since laws are calibrated to tell the truth only "in each separate domain" (Cartwright, 1983, p. 12).

3. Local Testability and a "Modular" Approach to Scientific Explanation

Ecology, experimental economics, psychology, biology, and the social sciences operate within open systems, where interactions are nonlinear, historical, and context-dependent. Testability in these sciences cannot follow the Newtonian physics model; instead, it presupposes locally situated testing, shaped by environment, population, historical period, and so forth.

Sandra Mitchell examines the difficulty of identifying universal laws and explores the situation of contingent-contextual laws in biology and the social sciences, introducing the concept of "modular explanation" (2000). Here she advocates for an epistemology of complexity, in which laws are contextual, contingent, and not strictly replicable. Mitchell envisions a "modular" approach to scientific explanation, in which laws operate locally through the combination of partial models, depending on purposes and instruments. Testing thus becomes a contextual practice, influenced by material, institutional, and technological factors.

Mitchell revisits the dispute between reductionism and pluralism, rejecting both unifying monism and the "anything goes "alternative (2003). She advances a pragmatic conception of laws: the laws useful to science are rarely universal but rather are instruments grounded in idealizations, useful for prediction, explanation, and intervention. This perspective does not exclude strong regularities but broadens the space of law to include different types of statements with varying degrees of generality and robustness, compatible with critical pluralism. Consequently, many apparently irreconcilable conflicts are, in fact, conflicts of representation rather than of substance. Models isolating partial causes can be internally consistent, while tension reemerges only when they are applied in an integrated manner to concrete cases, where causes interact.

4. Testing as Fit and Performance

The era of Big Data and computational simulations marks a shift in how scientific theories are formulated and tested. Traditionally, a theory was a collection of analytic propositions or mathematical models, testable through deduction and experiment. Today, in many fields such as ecology, genomics, computational physics, and behavioural economics, theories may take the form of algorithms, numerical models, or data-extracted patterns, often without explicit formal expression.

In climatology, quantum physics, or epidemiology, computational simulations have become epistemological instruments. The behaviour of such phenomena cannot be reduced to formulas, and theories are tested indirectly. A model is validated through its ability to reproduce observed data or the behaviours of complex systems, such as pandemic modelling, weather prediction, or conflict simulations. Here, the focus shifts from verifiability to calibration, parameter tuning, and cross-model comparison.

In the algorithmic context, the central question arises, how do we test a theory if it is formulated not through equations but as a machine learning algorithm? In such cases, precise prediction is replaced by probabilistic approximations of system behaviour. Models learn from datasets divided into training, validation, and test sets. Their performance is then evaluated using quantitative metrics such as accuracy or F1-score, which measure classification or prediction capacity. Accuracy represents the proportion of correct or true predictions out of the total made; the F1-score, the harmonic mean of precision and recall, is particularly useful in cases with imbalanced classes, where certain phenomena are rarer. These metrics reflect the empirical performance of the model, not necessarily its explanatory value or the truth of the theory.

Thus, an epistemological shift occurs: from testability as theoretical demarcation to testability as statistical optimization within algorithmic parameter space. In classical sciences, testing required formulating a hypothesis, conducting a controlled experiment, and confirming or refuting the hypothesis on the basis of data. In machine learning and data science, prediction becomes the primary goal, not explanation. Here, hypotheses are not explicitly formulated, and models are evaluated by how well they perform on test datasets. The result is an epistemology grounded in performance rather than truth.

This paradigm raises new issues, such as algorithmic opacity (black box situations), since deep learning models can deliver accurate results without causal explanations. Agent-based models, deep learning, and machine learning, grounded in radical inductivism, risk rendering theory irrelevant by discovering laws solely through pattern recognition. This creates the danger of overconfidence in empirical regularities without theoretical justification, as in correlations without causality.

In the age of Big Data, the logical structures of theories recede to the background, and testability becomes a function of data-processing infrastructure. Simulations and algorithms regenerate possible worlds, which we compare with reality through fit, relinquishing laws and deductive systems. This transformation calls for a reformulation of criteria of scientific validity, requiring robustness, transparency, reproducibility, and interpretability.

5. Data-Centric Testing

In contemporary sciences, data have become the epistemological infrastructure, surpassing their traditional role as mere empirical support for laws. Sabina Leonelli (2016) argues that testing has shifted from the verification of laws or hypotheses to the evaluation of patterns derived from massive datasets. She introduces the concept of distributed testability, according to which a theory is tested through multiple experiments, across networks of data, infrastructures, and computational flows.

Leonelli advocates replacing the term "data-driven" with a "data-centric" perspective, in which practices of data mobilization, integration, and visualization are recognized as acts of discovery in themselves, rather than as byproducts of theory testing. She observes the rise of a data-centric approach in which data-related efforts are valued as contributions to discovery in their own right. According to Leonelli, the convergence of digital technologies (for data production, dissemination, and analysis) with new institutional regulatory regimes has reordered research priorities, with significant effects on what we call "scientific knowledge" and on how such knowledge is produced, legitimized, and used.

Epistemic labour devoted to data collection, classification, and interpretation has become central. Leonelli's book empirically examines how online databases mediate the aggregation, mobilization, and evaluation of knowledge in the life sciences. She dismantles traditional theoretical centrism, the idea that only theories and explanations are prestigious research outcomes, by showing that data-handling practices have long been undervalued and relegated to technical and archival staff, often considered mere support personnel and rarely recognized as contributors to knowledge production.

Leonelli situates this reorientation within a broader context, shaped by high-throughput technologies (such as NGS sequencing and microarrays) and open science policies, which have instilled a "data ethics" into the institutions of science. Editorial rhetoric captures this shift with the expression "If you have useful data forgotten in a drawer, send them for publication" (Leonelli, 2016, p. 4).

The material, social, and institutional pathways through which data are "packaged" and transported in order to function in new contexts contribute to their relational character, what counts as data depends on their evidential value in a given research situation, highlighting the role of experimental know-how and classificatory theories. Data-centrism brings to the forefront the infrastructures, procedures, and institutions that enhance the value of data through standardization, openness, curation, governance, and interoperability.

Philosophically, data-centrism fosters a relational theory, in which data acquire meaning through the ways in which they are mobilized within scientific and technical networks. Underdetermination becomes an epistemic engine, since the same dataset may support multiple explanations. The practical implication is clear: emphasis shifts from mere data accumulation toward investments in epistemic infrastructure, through careful curation, well-structured metadata, intelligent sharing policies, and reflexivity regarding biases, responsibilities, and implicit limitations.

This orientation makes data-centrism both compatible with current political and economic regimes and a fertile ground for philosophical and sociological critique of contemporary science.

6. From Replicability to Contextual and Extra-Empirical Theoretical Confirmations

Traditionally, replicability was the primary criterion for testing theories. If an experiment could be repeated under the same conditions, the theory was thereby supported. In contemporary sciences, however, data are not always public or replicable. Algorithms have become sensitive to parameter tuning, training sets, and even the version of the software libraries employed. Today, testing is increasingly tied to the institutionalization of the very idea of testing, to a form of scientific validation without universality. Against a backdrop that shifts from falsification to situated verification, replicability is contrasted with robustness in algorithmic science. A theory is considered valid because it produces consistent results across varied contexts, without needing to be identically replicated. A special case of theoretical testing is provided by present-day cosmology and physics, where experimental confirmations are, in practice, no longer possible.

Richard Dawid (2013) traces the trajectory of the last decades, which fulfilled in remarkable ways the expectations physicists had in the 1970s. The Standard Model was indeed experimentally confirmed and led to an extraordinary sequence of predictive successes. The most recent Standard Model prediction found empirical confirmation in the summer of 2012 through the Large Hadron Collider (LHC) experiments at CERN, thereby closing an important phase in the evolution of fundamental physics.

Moreover, theoretical progress has continued well beyond the Standard Model, giving rise to a series of ambitious new frameworks. Grand Unified Theories (GUTs) postulate a more unified structure of nuclear interactions. Supersymmetry (SUSY) proposes an extended symmetry connecting particles of different spin. Supergravity further expands this idea toward a general theory of gravitation. String theory emerged as the first strong and promising candidate for a unified theory of all physical interactions. Finally, inflationary cosmology substantially reshaped our perspective on the early universe, bringing cosmological model-building into closer alignment with high-energy physics.

Yet, all the theories mentioned that go beyond the Standard Model share a common problem. Although each was formulated several decades ago, none has received empirical confirmation to date. The canonical experimental strategy of testing at increasingly higher energies, by constructing ever-larger particle accelerators, has become exceedingly difficult to sustain, due to the massive efforts required to increase energy levels. The giant LHC experiment at CERN may well be the last of its kind. Supersymmetry remains the only major theory beyond the Standard Model that still has a realistic chance of empirical confirmation within LHC experiments. All the other theories require characteristic energy levels that are expected to lie far beyond any technically feasible parameters.

Cosmic inflation shows a degree of promising consistency with cosmological data, but its foundational principles remain exceedingly difficult to confirm conclusively. The other theories have little hope of achieving significant empirical confirmation in the foreseeable future. The increasing detachment of theoretical construction from empirical confirmation may be regarded as heralding the dawn of a profound crisis in fundamental physics.

7. A Modified Perspective on Scientific Realism

Following his analysis of physical theories, Richard Dawid (2013) evaluates a broader philosophical issue, to what extent does science provide knowledge about the reality of the world, beyond the organization of observations and empirical predictions? This is the classical problem of the debate on scientific realism, in which realists maintain that theories disclose truths about the unseen world, while anti-realists argue that they remain merely instruments for ordering phenomena.

To contextualize the issue, Dawid revisits the longstanding dispute between rationalism and empiricism. Rationalism claimed that pure reason could generate stable knowledge about the world, whereas empiricism insisted that genuine knowledge must derive from observation. The success of modern scientific methodology consecrated the supremacy of empiricism and fuelled scepticism toward rationalist speculation. But if science relies exclusively on empirical data, can we say that it describes anything more than those data themselves? The 19th-century debate over atomism clearly illustrated this tension. Realists argued that the empirical success of the atomic hypothesis demonstrated the existence of atoms, while anti-realists regarded it as a useful but ontologically unjustified construction. Later, 20th-century physics discoveries tilted the balance toward a pragmatic use of microphysical concepts, leaving the ontological problem largely to philosophers.

7.1. The No Miracles Argument and the Unconceived Alternatives Argument

A classical argument for scientific realism is the "no miracles argument", formulated by Hilary Putnam (1975, 1978), according to which it would be inexplicable, indeed "miraculous", for scientific theories to achieve repeated predictive success if they were not at least approximately true. This type of reasoning, as an inference to the best explanation, provides realism with an empiricist-inductive justification.

Anti-realists have responded with meta-inductive scepticism, arguing that the history of science offers numerous examples of theories that were both predictively successful and ontologically false. Along this line lies the "unconceived alternatives argument", explicitly formulated by Kyle Stanford (2006) and prefigured by Lawrence Sklar (1975), which holds that for any current theory there most likely exist viable alternatives not yet conceived. Therefore, we cannot realistically justify the conviction that present theoretical structures faithfully refer to reality. This critique highlights that realist claims rest less on direct empirical confirmation than on estimations of the degree of scientific underdetermination.

This is also Dawid's interpretation, evaluations of the limitations of underdetermination can account for the predictive success of science without presupposing realism. Instead of asserting that theories are approximately true, one may claim that their success derives from the rarity of coherent theories that fit the data. This explanation is more flexible than realism and more resistant to historical objections raised by pessimistic meta-induction.

Dawid (2013) introduces an additional element. By presenting itself as a "final theory," it suggests that the limits of underdetermination may be so stringent that no viable alternatives exist. In such a case, realism would gain stronger support, insofar as the theory could represent the only consistent structure compatible with empirical data. Yet, string dualities challenge classical ontological realism, as they demonstrate that entities with very different characteristics can describe the same phenomenology. Faced with this situation, the only viable realist stance is to focus on structures rather than objects.

This is what Dawid terms "consistent structure realism". This position rests on the idea that reality is not constituted by stable ontological objects but by unique, coherent, and rare scientific structures, whose existence is motivated by the strong constraints imposed by underdetermination. Science, in this view, pursues truth about the world by pushing the boundaries of underdetermination, a strategy that extends empiricism without abandoning rationalist speculation.

7.2. Science, technology, fiction

Bruno Latour looks at science and technology "in action" (1987), that is, in the living process of research, before results are fixed as indisputable truths. Within this framework, the notion of the "black box" plays an epistemic role. A black box is a device, a theory, or a scientific result whose internal functioning is no longer questioned. It is accepted as such and used as an instrument for further work. When we say that DNA has the form of a double helix, we no longer return to Watson and Crick's demonstrations; we simply operate with that fact. In the same way, we use the computer without worrying about the electronic circuits that make it work.

The distinction emphasized by Latour is that between "established" science and "science in action." The former refers to sedimented knowledge, where facts and theories are treated as self-evident data, no longer subject to controversy. The latter designates the process through which hypotheses, experiments, and prototypes are fragile, disputed, and open to criticism. By studying this second dimension, Latour shows that we can better understand how scientific knowledge is actually produced.

A crucial point in his argument is that facts and artifacts do not simply emerge naturally, nor do they impose themselves by virtue of an intrinsic essence; they are the result of collective processes of research, negotiation, and stabilization. There is no clear line between the discovery of a fact and the construction of a technological object. Both involve the writing of texts, the performance of experiments, the use of instruments, the mobilization of institutions, and, ultimately, the achievement of consensus. Latour sets out to "open the black boxes" and to trace the steps by which a hypothesis becomes an accepted fact or an uncertain prototype is transformed into a robust technology. This transformation involves recruiting allies (other researchers, funders, institutions) building support networks, elaborating arguments and experiments, and gradually consolidating results until they appear inevitable and objective. The power of science, Latour argues, lies in its capacity to mobilize dense networks of people, instruments, texts, and organizations.

This perspective has significant consequences. To analyse science in action is to render it more intelligible. It becomes visible why some theories succeed while others are abandoned, why certain technologies take hold while others disappear, and how knowledge remains constantly connected with the economy, politics, and culture. Science appears not as a set of absolute truths, but as a dynamic social practice, a field in which people create, negotiate, test, and stabilize facts and artifacts. Once consolidated, these results themselves become the starting point for new cycles of research and innovation.

The social impact of science, and the process by which it arises, is strikingly similar to that of literary fiction, as Latour also notes (2005). The apparent distance between the two diminishes once we carefully observe the conditions under which a scientific hypothesis comes to be accepted or rejected. In both science and literature, everything begins with a fragile construction that must capture attention, persuade, and circulate effectively in a social environment. Scientific statements take shape together with a whole network of texts, images, instruments, graphs, and experimental protocols. In both cases, what is at stake is the telling of a story. The difference lies in the way the story acquires consistency and becomes anchored in instruments, contexts, and institutions.

Latour stresses that the robustness of a fact derives less from the statements it makes than from the material trajectory it follows. Inscriptions, data, graphs, diagrams, and instruments all contribute to consolidating a claim and transforming it into a reality that is difficult to dispute. Stability thus appears as the outcome of a distributed and cumulative process rather than as a reality in itself. It is this collective work that marks the difference between a hypothesis that remains marginal and one that becomes a reference point in the scientific field.

For sociologists and researchers in the social sciences, this perspective reshapes the traditional approach. Instead of assuming that scientific facts are immutable givens, clearly demarcated from the domain of fiction, it is more relevant to examine how they are constructed, who supports them, by what means, in what institutional contexts, and through what channels they circulate. The aim is not only to recognize the value of scientific knowledge, but to understand what makes its solidity possible.

The absence of a clear ontological boundary between what we call “fact” and what we consider “fiction” means that the difference between them depends less on the nature of the narrative than on the degree of collective, infrastructural, and institutional investment from which they benefit. Facts thereby acquire a relative degree of “reality,” not by virtue of a hidden essence, but because they endure within a complex ecosystem of interactions and supports of a social kind.

8. Testing as an Institutional and Social Practice

What is considered testable increasingly depends on funding agencies and projects that demand “clear” testability, with priority often given to fields such as medical AI or climatology. Scientific journals typically require standardized, reproducible methods, which are often incompatible with exploratory practices. Disciplinary norms differ. For instance, in theoretical physics indirect testing is acceptable, whereas in psychology it is not. Thus, testing is transformed into a socio-technological practice, mediated by instruments, norms, and collective interests.

The fact that testing is no longer primarily an epistemic or theoretical practice brings with it certain risks and epistemological tensions. When testing focuses solely on empirical performance without a theoretical model, there is the risk of losing explanatory knowledge. Machine learning, in particular, can lead to confusion between correlation and causation, especially in situations where testability is captured by technological infrastructure. Under such conditions, what counts as testable becomes restricted to what can be computationally modelled.

9. Toward an Adapted Philosophy of Testability. Methodological Pluralism and Infrastructures of Validation

In contemporary science, it is no longer possible to speak of a single form of testability valid across all disciplines. Each field operates with its own criteria of validation, tailored to its theoretical and methodological specificities. In biology, ecological robustness and resistance to variations in experimental conditions serve as signals of epistemic reliability. In artificial intelligence, validation occurs through performance on datasets and, increasingly, through algorithmic interpretability criteria. In physics, testing is often based on indirect confirmations via observable effects, as in the case of the detection of the Higgs boson.

These differences do not imply relativism, but rather the necessity of a methodological pluralism that acknowledges the legitimate diversity of testing practices. Such pluralism grounds scientific coherence in the epistemic realities of different domains. In this context, the robustness of a theory could be measured through the design of a multidimensional evaluation matrix, including factors such as stability across varied datasets, sensitivity to initial conditions, conceptual transparency, and operational reproducibility. Composite robustness indices could be developed, analogous to statistical validity but extended computationally and epistemologically.

This approach would make possible not only cross-theory comparison but also differentiated evaluation between testing in the formal sciences (mathematics, logic), where validation is deductive and internal, and in the empirical sciences, where testing is contingent, experimental, or algorithmic. A comparative philosophy of validation would function as a critical map of testing norms and epistemic rationalities across the regions of science.

In this way, theory testing may become an interdisciplinary space where philosophy, the sociology of science, and computational sciences must cooperate. Integrating these domains enables a more comprehensive analysis of how theories are validated, logically, empirically, institutionally, technologically, and ethically.

Science policies, from grants and peer review mechanisms to requirements for impact or "deliverables", often shape what is considered testable. Funding agencies, for example, tend to privilege "visible" and measurable testability at the expense of speculative, theoretical, or interdisciplinary research. A critical approach must therefore expose and challenge these validation norms imposed from outside the epistemic process itself.

It is also essential to extend testing to models that cannot be reduced to equations or deductive statements. Simulations, agent-based models, neural networks, and machine learning algorithms bring alternative forms of validation, grounded in performance, consistency, and computational transparency. In fields such as astrophysics or genomics, theories are tested indirectly, through emergent effects or statistical patterns. In artificial intelligence, especially in deep learning, hypothesis testing is replaced by the selection of models according to their performance.

All these transformations indicate that testability has become a complex social practice with material, institutional, and technological dimensions. Scientific laws, once considered universal and abstract, are now treated as local and context-dependent. This complex landscape calls for a flexible philosophy of testability, one that integrates methodological pluralism, critical analysis of the social conditions of research, and reflection on the infrastructures that make the validation of a theory possible, or impossible.

References

Cartwright, N. (1983). *How the laws of physics lie*. Oxford University Press.

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

Leonelli, S. (2016). *Data-centric biology: A philosophical study*. University of Chicago Press.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Latour, B. (1987). *Science in Action.* Harvard University Press.

Mitchell, S. D. (2000). *Dimensions of scientific law*. Cambridge University Press.

Mitchell, S. D. (2003). *Biological complexity and integrative pluralism*. Cambridge University Press.

Putnam, H. (1975). *Mathematics, matter and method: Philosophical papers, Vol. 1*. Cambridge University Press.

Putnam, H. (1978). *Meaning and the moral sciences*. Routledge & Kegan Paul.

Sklar, L. (1975). Methodological essentialism, the interpretive models of empirical science, and the interpretive adequacy of theories. *Philosophy of Science, 42*(3), 331–348. https://doi.org/10.1086/288600.

Stanford, K. (2006). *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. Oxford University Press.

van Fraassen, B. C. (1980). *The scientific image*. Oxford University Press.