PRINCIPIA **30**(3): 655–672 (2026) doi: 10.5007/1808-1711.2026.e109429
Published by NEL —Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

ON THE EFFICACY OF INFERENCE TO THE BEST EXPLANATION IN THE METAPHYSICS OF SCIENCE

ROBERTO MIGUEL AZAR

Universidad Argentina de la Empresa (UADE), ARGENTINA
University of Buenos Aires (UBA), ARGENTINA
robertoazar86@gmail.com
https://orcid.org/0000-0002-5751-5510

BRUNO BORGE

University of Buenos Aires (UBA) / CONICET, ARGENTINA
brunojborge@gmail.com
https://orcid.org/0000-0002-1755-9690

GABRIEL CHIAROTTI SARDI

Universidade Estadual do Centro-Oeste do Paraná (UNICENTRO), BRAZIL
gabrielsardi@unicentro.br
https://orcid.org/0000-0002-6320-0400

Abstract. This paper examines the role of Inference to the Best Explanation (IBE) in the metaphysics of science, a field concerned with understanding the fundamental nature of reality through engagement with scientific theories and models. While IBE is commonly used to justify belief in scientific theories, its application in metaphysics is debated. Some philosophers, like Armstrong, Ellis, and Swoyer, argue that IBE can also legitimize metaphysical commitments due to their explanatory power. However, critics like Saatsi challenge this parallel, questioning the reliability of IBE even within science, given its mixed historical record. To address this debate, the paper draws on a distinction proposed by Azar and Sardi between two senses of IBE. The first, creative sense—akin to abduction—involves generating novel explanatory hypotheses and is considered valuable for metaphysical inquiry when understood as an exploration of possibilities informed by science. The second, selective sense—choosing the best explanation among known alternatives—faces significant challenges in metaphysical contexts, as its truth-conduciveness has been questioned by several authors (Saatsi 2017, Allzén 2023, Ritchie 2022) due to the absence of the empirical feedback mechanisms that support its reliability in science. We argue that, regardless of how the debate over the epistemic status of selective IBE in metaphysics is ultimately resolved, a moderately naturalized metaphysics of science can benefit from abduction while retaining its explanationist commitment.

Keywords: inference to the best explanation • abduction • metaphysics of science • moderately naturalized metaphysics

RECEIVED: 21/10/2025 REVISED: 30/03/2026 ACCEPTED: 18/09/2026 PUBLISHED: 30/10/2026


© 2026 The author(s). Open access under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License.

1. Introduction

There are ongoing debates concerning the efficacy of employing Inference to the Best Explanation (IBE) as a reliable method in both scientific and philosophical inquiry. While IBE is often invoked to justify belief in scientific theories, its applicability and effectiveness within the domain of metaphysics remain a contentious issue. This paper specifically examines the efficacy of IBE within the metaphysics of science, an area that seeks to understand the fundamental nature of the world by engaging with the outputs of scientific theories, models, and practices.

Proponents of explanationism in metaphysics, such as Armstrong (1983, 1997), Ellis (2009), and Swoyer (2008), often draw parallels between metaphysical theorizing and the explanatory practices of science, suggesting that IBE can similarly justify belief in metaphysical entities and principles by virtue of their explanatory power and unifying capabilities. Armstrong, for instance, likens his argument for laws of nature to explanatory inferences to theoretical entities in science (1983, p.104). On the contrary, Saatsi (2017) is skeptical towards this straightforward application. He highlights the challenges in establishing a robust methodological continuity between science and metaphysics regarding IBE, particularly concerning the qualified reliability of IBE in science itself given its historical track record. This skeptical line has been further developed by other authors. Allzén (2023) argues that the reliability of IBE in science crucially depends on empirical feedback loops, predictive testing, and systematic error correction—conditions that are paradigmatically absent in metaphysical contexts—and that, consequently, the inference from the reliability of abductive methodology in science to its reliability in metaphysics is unwarranted. Ritchie (2022) develops a related critique, maintaining that explanatory reasoning and appeals to theoretical virtues such as simplicity have no more than a formal analogue in metaphysics: truth-conduciveness in the sciences does not entail truth-conduciveness in metaphysical inquiry. Bryant (2017), while sharing some of these epistemic concerns, adopts a more moderate stance: she argues that what she calls "free range metaphysics" is epistemically inadequate insofar as it relies on criteria such as consistency, simplicity, and explanatory power without empirical grounding, but she nonetheless recognizes that it plays valuable collateral roles, including conceptual clarification and the incubation of ideas that may prove fruitful for further inquiry. On the other side of the debate, several authors defend methodological continuity between science and metaphysics. Paul (2012) conceives of metaphysics as a form of model-building continuous with scientific practice, arguing that both domains employ similar strategies of theory construction and evaluation guided by theoretical virtues and inference to the best explanation. Williamson (2016) extends this claim further, defending a broadly abductive methodology shared across philosophy and science, on which philosophical theories—including metaphysical ones—are selected

by means of the same kind of explanatory reasoning operative in the sciences. The epistemic status of IBE in metaphysics thus remains an open and actively disputed question in the contemporary literature.

In this context, our proposal is to consider a distinction made by Azar (2017, 2019) and Sardi (2022) between two senses of IBE. The first sense corresponds to a creative, ampliative step, akin to abduction, which involves generating novel explanatory hypotheses. The second sense refers to a selective step, where the best explanation is chosen from a set of pre-existing alternatives. We argue that the first, creative sense of IBE, understood as abduction, holds significant utility for the metaphysics of science when the latter is conceived as an exploration of the realm of possibility (Morganti & Tahko 2017). This involves investigating metaphysical questions that arise from or are informed by our best scientific theories and models, essentially exploring the space of metaphysically plausible scenarios consistent with scientific understanding. Conversely, we argue that there are substantial reasons to question whether the second, selective sense of IBE can be reliably transferred to metaphysical contexts. As several authors have argued (Saatsi 2017, Allzén 2023, Ritchie 2022), the truth-conduciveness of selective IBE may depend crucially on empirical feedback mechanisms that are absent in metaphysics. While the extent of this limitation remains debated, we contend that the abductive, creative sense of IBE provides a clearly productive role for metaphysical inquiry regardless of how this broader dispute is resolved.

While the individual elements of this argument have been developed separately in the existing literature—the distinction between abduction and IBE as inferential processes (Azar 2017, 2019; Sardi 2022; Sardi & Silva 2023), the debate over the reliability of IBE in metaphysical contexts (Saatsi 2017; Allzén 2023; Ritchie 2022), and the proposal of a moderately naturalized metaphysics of science (Morganti & Tahko 2017)—to the best of our knowledge, no systematic integration of these threads has been offered. The present paper provides such an integration by showing how the conceptual distinction between abduction and selective IBE directly addresses the challenges raised against explanationism in the metaphysics of science, and it does so while remaining agnostic with respect to the broader debate on methodological continuity between science and metaphysics.

2. Moderately Naturalistic Metaphysics

The emergence of metaphysics of science as a distinct philosophical discipline represents one of the most significant developments in contemporary analytic philosophy. This field arose from the convergence of two complementary movements: the growing interest among philosophers of science in metaphysical questions about the

nature of scientific entities and structures, and the increasing attention paid by metaphysicians to the implications of our best scientific theories for fundamental ontological questions.

Historically, the relationship between metaphysics and science has been marked by periods of both collaboration and tension. The logical positivists of the early twentieth century famously attempted to eliminate metaphysics altogether, viewing it as meaningless speculation that stood in opposition to the empirical rigor of science. However, the subsequent recognition that scientific theories inevitably involve metaphysical commitments—about the nature of space and time, causation, natural kinds, and theoretical entities—led to a gradual rehabilitation of metaphysical inquiry within the philosophy of science community.

The contemporary field of metaphysics of science emerged in response to this recognition, seeking to develop systematic accounts of the metaphysical foundations and implications of scientific theories. As Soto (2015) notes, this discipline is characterized by two primary objectives: producing metaphysical proposals that are scientifically informed, and elucidating scientific concepts through the tools of metaphysical analysis. This bidirectional relationship distinguishes metaphysics of science from both traditional a priori metaphysics and purely empirical scientific investigation.

The methodological foundations of metaphysics of science remain a subject of considerable debate. At the heart of this debate lies a fundamental question about the relationship between metaphysical and scientific methods: to what extent should metaphysical inquiry be constrained by, or continuous with, the methods employed in empirical science? This question has profound implications for how we understand the epistemic status of metaphysical claims and the criteria by which such claims should be evaluated.

Proponents of strong continuity argue that metaphysical hypotheses should be subject to the same standards of empirical accountability as scientific theories, or at least should be directly answerable to empirical considerations. On this view, metaphysical claims derive their legitimacy primarily from their ability to contribute to our understanding of the empirical world, whether through explanation, prediction, or theoretical unification. This perspective emphasizes the overlap between metaphysical and scientific subject matter—both disciplines, it is argued, are fundamentally concerned with understanding the nature of reality as revealed through empirical investigation.

Conversely, advocates of methodological discontinuity maintain that metaphysics possesses its own distinctive methods and epistemic standards that, while potentially informed by scientific findings, are not reducible to empirical testing. They argue that metaphysical inquiry operates primarily through a priori reasoning about conceptual possibilities and necessities, and that its value lies not in empirical predictive power but in providing conceptual frameworks for understanding the fundamental struc-

ture of reality. This view emphasizes the autonomy of metaphysical reflection and its capacity to explore possibilities that may extend beyond what is directly accessible through empirical investigation.

The debate over continuity versus discontinuity in methods is closely connected to questions about the overlap of subject matter between metaphysics and science. While some philosophers argue that both disciplines are concerned with the same fundamental aspects of reality—the nature of space, time, causation, and material objects—others contend that metaphysics deals with more abstract or general features that are conceptually prior to the specific entities and processes studied by particular sciences. This difference in emphasis leads to different views about whether metaphysical and scientific inquiries can be pursued independently or require systematic integration.

In the present context, these debates about methodological and substantive continuity are particularly relevant because our investigation concerns the legitimacy of applying Inference to the Best Explanation—a method widely recognized as characteristic of scientific practice—within the domain of metaphysics of science. This question is at the heart of contemporary debates about explanationism in metaphysics—the methodological approach that places explanatory considerations at the center of theoretical justification and draws heavily on IBE for supporting metaphysical claims. Explanationists like Armstrong (1983, 1997), Ellis (2009), and Swoyer (2008) argue for methodological continuity between science and metaphysics, suggesting that just as IBE can justify belief in unobservable scientific entities, it can similarly warrant commitment to metaphysical posits by virtue of their explanatory power.

This explanationist stance finds broader support in the literature on methodological continuity. Paul (2012) conceives of metaphysics as model-building continuous with scientific practice, where metaphysical theories are developed, compared, and selected using IBE and theoretical virtues in a manner structurally analogous to scientific theory choice. Williamson (2016) defends an even more encompassing view, arguing that abductive methodology is shared across philosophy and science, with no principled methodological boundary separating the two.

However, this explanationist program faces significant challenges. Saatsi (2017) has argued that the absence of empirical feedback mechanisms in metaphysics undermines the reliability that IBE enjoys in scientific contexts. This concern has been reinforced by Allzén (2023), who contends that the conditions that make IBE reliable in science—empirical feedback, predictive testing, and systematic error correction—are absent in metaphysical contexts, rendering the inference from scientific to metaphysical reliability unwarranted. Ritchie (2022) adds that appeals to theoretical virtues such as simplicity possess no more than a formal analogue in metaphysics, and that their truth-conduciveness in the sciences cannot be straightforwardly extended to

metaphysical inquiry. From a different angle, Chakravartty (2017) proposes understanding the relationship between science and metaphysics in terms of a continuum of epistemic risk, suggesting that there is no privileged point at which to draw the line between empirically warranted and epistemically risky metaphysical claims.

We adopt a neutral stance toward these debates, recognizing that both continuity and discontinuity positions capture important insights about the nature of metaphysical inquiry. Rather than attempting to resolve these fundamental methodological questions, our aim is to examine how the efficacy of IBE in metaphysics of science might be understood regardless of one's position on these broader issues. This approach allows us to focus on the specific features of IBE that are relevant to metaphysical reasoning while remaining open to different conceptions of metaphysics' relationship to empirical science.

The contemporary landscape of metaphysics of science thus presents us with a complex picture: a discipline that has established its legitimacy through engagement with scientific theories and practices, yet continues to grapple with fundamental questions about its own methods and epistemic foundations. It is within this context that questions about the role of explanatory reasoning in metaphysical inference become particularly pressing, as they touch on core issues about the nature of metaphysical knowledge and its relationship to empirical investigation.

3. Distinction between Abduction and Inference to the Best Explanation

The terms '*abduction*' and '*Inference to the Best Explanation*' (IBE) are often treated as synonyms in the philosophy of science. However, this identification overlooks an important conceptual and methodological distinction, which has significant implications when assessing the reliability and scope of explanationist strategies in metaphysical inquiry.

Historically, the notion of abduction originates with Charles S. Peirce (1958), who characterized it as an ampliative mode of inference, distinct from deduction and induction. For Peirce, abduction is the process by which a hypothesis is generated that could plausibly explain surprising or puzzling facts. It is therefore fundamentally creative: a leap from observed phenomena to possible explanatory causes. As Peirce famously noted, "Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea" (CP 5.171).

That is, abduction, according to Peirce, is a creative process influenced by the researcher's background knowledge and the context in which they operate. It does not begin with a comparison among pre-existing hypotheses. Instead, it works as an instinctive (CP 5.173) and fallible (EP2, p.227) insight, constituting the starting point

of scientific inquiry: the moment when something new is formulated.[1]

> The abductive suggestion comes to us like a flash. It is an act of insight, although of extremely fallible insight. It is true that the different elements of the hypothesis were in our minds before; but it is the idea of putting together what we had never before dreamed of putting together which flashes the new suggestion before our contemplation (EP2, p.227).

For example, someone is engaged with the problem of how to proceed in constructing a typewriter. After a prolonged period of reflection, drawing upon their background knowledge related to this issue, various ideas remain diffusely present in their mind, none of which, when taken individually, bears a direct relation to the central problem. At a certain point, however, these ideas begin to articulate themselves in a novel way, forming a combination that provides a possible explanation for the difficulty at hand and that, for this reason, suddenly emerges into consciousness with great vividness (CP 7.498). This constitutes a paradigmatic case of creative abduction, insofar as the solution does not arise from a prior logical sequence or from an already established habit, but rather from an innovative synthesis that imposes itself by virtue of its explanatory force.[2]

In contrast, Gilbert Harman's classic account (1965) reformulated the idea under the label 'Inference to the Best Explanation'. For Harman, IBE is the inferential strategy by which one selects, from among a set of pre-existing hypotheses, the one that best explains the available evidence. This conception shifts the focus from generation to selection: IBE does not introduce new explanatory candidates but presupposes a set of alternatives and provides a criterion for preferring one over the others.[3]

> In making this inference one infers, from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis. In general, there will be several hypotheses which might explain the evidence, so one must be able to reject all such alternative hypotheses before one is warranted in making the inference. Thus one infers, from the premise that a given hypothesis would provide a "better" explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true (Harman 1965, p.89).

Paul Thagard (1978) reinforced the idea that IBE consists in a process of comparative selection among rival alternative hypotheses—rather than in the generation of new alternatives—emphasizing that the choice of the best explanation must be guided by specific criteria of theoretical evaluation, often referred to in the philosophy of science literature as *cognitive values*, *epistemic values*, or *explanatory virtues*.

> To put it briefly, inference to the best explanation consists in accepting a hypothesis on the grounds that it provides a better explanation of the evidence

> than is provided by alternative hypotheses. We argue for a hypothesis or theory by arguing that it is the best explanation of the evidence (Thagard 1978, p.77).

Despite Peter Lipton's well-known reformulation of IBE (2004), in which he sought to combine the notions of hypothesis generation and selection through two distinct filters, Lipton still retained as the central element the idea of a comparative evaluation of alternatives based on scientists' background knowledge and the explanatory virtues of the hypotheses—such as simplicity, fertility, unification, and so on (Sardi 2024). This does not align with the core notion of creative insight inherent in Peircean abduction, since the comparative analysis of explanatory virtues goes beyond that inferential dimension.

> According to Inference to the Best Explanation, our inferential practices are governed by explanatory considerations. Given our data and our background beliefs, we infer what would, if true, provide the best of the competing explanations we can generate of those data (so long as the best is good enough for us to make any inference at all). Far from explanation only coming on the scene after the inferential work is done, the core idea of Inference to the Best Explanation is that explanatory considerations are a guide to inference (Lipton 2004, p.56).

In fact, Lipton attributed importance to the generative process of hypotheses in his first filter, which would function as a preliminary stage in which only genuinely plausible explanatory alternatives are selected or generated (Lipton 2004, p.59). This first filter would be guided by the context of the problem to be explained, the available data, and the influence of explanatory virtues—which would not be subjected to rigorous scrutiny (as occurs in the second filter of his model) but would instead influence the subject almost unconsciously and automatically. In this sense, the first filter in Lipton's IBE can be interpreted as a variation of Peirce's abduction, insofar as it allows for hypothesis generation, but it differs by incorporating explanatory virtues into its conceptual framework. In any case, his model remains essentially distinct from abduction, whether subtly in the first stage or radically in the second.

In summary, the versions of IBE proposed by Harman, Thagard, and Lipton are evaluative and comparative selection processes that go beyond the instinctive generation of explanatory possibilities. They require, as central elements of the selection and justification process, both the empirical confirmation of the hypothesis and a careful analysis of its explanatory virtues.

This bifurcation—between creative generation and selective evaluation—has been widely overlooked in contemporary debates,[4] generating confusion about the kind of inferential work IBE is supposed to perform, especially when it is extended beyond empirical science. As has been argued in recent work (Azar 2017, Azar 2019,

Sardi 2022, Sardi & Silva 2023), conflating abduction with IBE misrepresents both their logical structure and their epistemic credentials.

To illustrate this difference, one can recall a classic example: Kekulé's dream of a snake biting its own tail, which led him to propose the cyclic structure of benzene. This moment exemplifies abduction in the Peircean sense: an imaginative leap that formulates a novel explanatory hypothesis that was not previously available for selection. By contrast, the subsequent process in which chemists compared Kekulé's ring structure with rival linear models exemplifies IBE in its selective sense: an evaluation of which candidate better explains experimental data, such as substitution patterns. Thus, abduction is indispensable for expanding the explanatory space, while IBE, in its stricter sense, is essential for limiting it. Each plays a distinct and complementary role in the scientific process: discovery versus justification.

Furthermore, beyond distinguishing it from selective IBE, it is important to emphasize that abduction, as conceived by Peirce, involves a creative spark that is nonetheless fallible. Generating innovative hypotheses does not guarantee their truth: every abductive inference remains subject to error and, in science, requires subsequent verification through inductive methods. As noted by Sardi (2022, p.70) and Campos (2011, p.428), although abduction operates as a generative insight that inaugurates a novel explanation, its outcomes are conditioned by fallibilism, since its truth value can only be inferred after being tested. This distinction, repeatedly emphasized by Peirce, shows that abduction occupies a specific place within the research process: it is part of the context of discovery, whereas induction and empirical testing belong to the context of justification.

This shared fallibility, however, does not entail that abduction and selective IBE face the same difficulties when transferred to metaphysical contexts. The crucial difference lies in their respective functional roles within inquiry. Selective IBE operates in the context of justification: it aims to establish which hypothesis, among preexisting alternatives, is most likely to be true. Its reliability in science depends on the availability of empirical feedback to confirm or disconfirm the selected hypothesis. When this feedback is absent, as in metaphysical contexts, the selective function loses the very mechanism on which its epistemic credentials depend. Abduction, by contrast, operates in the context of discovery: it generates novel hypotheses without claiming to establish their truth. Its value lies not in confirmation but in expanding the space of explanatory possibilities. This generative function does not depend on empirical feedback in the same way, because it does not aim to confirm but to explore. The fallibility of abduction is therefore not a weakness of the same kind: it is an inherent feature of a process whose purpose is to open new lines of inquiry, not to close them.[5]

Peirce's understanding of abduction finds resonance in Norwood Hanson's observations (2010, p.70–73), especially regarding the practices of physics. According to

Hanson, physicists do not begin their work by testing previously established hypotheses through induction. Rather, they start from the observation of concrete phenomena and, based on those data and already consolidated knowledge, they formulate explanatory hypotheses. The confirmation stage—which may occur through enumerative induction or other methods—appears only at a later point. Thus, for Hanson, the creative process of hypothesis formulation should not be confused with the domain of induction, nor with selective versions of IBE. It is in this context that Peirce's notion of abduction assumes a central role: it is the type of reasoning responsible for introducing new hypotheses, with explanatory ambitions though still provisional. Once generated, these hypotheses become candidates for empirical testing and critical evaluation by the scientific community.

This compatibility between the creativity of abduction and its fallibility supports what can be called a "reasonably moderate epistemic optimism" (Azar 2023): we trust in the human mind's capacity to generate plausible hypotheses and make progress in understanding reality, albeit always under the condition of their testing and potential refutation. As Soto (2014) reminds us, Peirce distinguishes between first-order abductions—those that postulate a general fit between mind and world—and second-order abductions, which are the concrete hypotheses arising in scientific practice. The famous discovery of Neptune by Le Verrier, the hypothesis of dark matter, or the postulation of dark energy are paradigmatic examples of second-order abductions that illustrate the fruitful, though fallible, role of this mode of inference in science.

Recognizing this framework allows us to understand why abduction, even when detached from the phase of immediate empirical justification, can support a fertile explanationism when applied to contexts such as the metaphysics of science: it offers the power to explore possible scenarios while simultaneously maintaining the caution that every hypothesis, no matter how innovative, remains open to revision and critique.

Thus, distinguishing between abduction and IBE clarifies their logical structure and, more importantly, reveals that they face different challenges when applied outside the domain of empirical science. While neither is immune to error and both ultimately require some form of corroboration in scientific contexts, they occupy distinct positions within the research process—discovery and justification, respectively—and this difference has significant consequences for assessing their potential contributions to metaphysical inquiry.

4. IBE in the metaphysics of science

IBE has become a central tool for legitimizing beliefs in scientific contexts. Philosophers such as Armstrong (1983, 1997), Ellis (2009), and Swoyer (2008) have ex-

tended its scope to the realm of metaphysics, arguing that, just as IBE supports theories about unobservable entities in science, it can similarly justify positions concerning the fundamental structure of reality. On this view, the postulation of metaphysical entities—such as universal properties, dispositions, or laws of nature—can be justified by virtue of their explanatory power and unifying capacity. Armstrong, for example, explicitly compares his argument in favor of laws of nature as relations between universals to the explanatory inferences that posit theoretical entities in physics or chemistry (1983, p.104), suggesting that metaphysics shares the same methodological legitimacy when employing IBE. As noted in Section 2, this explanationist stance finds broader support in the work of Paul (2012) and Williamson (2016), who defend methodological continuity between science and metaphysics on the basis of shared abductive practices.

However, this transposition of IBE from science to metaphysics faces a central difficulty: the absence of mechanisms for empirical control. In the scientific realm, selective IBE relies on the possibility of subjecting rival hypotheses to experimental or observational testing. Coherence with new data, predictive capacity, and the possibility of falsification constitute much of the epistemic support that bolsters its credibility. As Saatsi (2017) argues, this support may be significantly diminished when IBE is transferred to metaphysics, since metaphysical theories typically lack testable empirical consequences. Allzén (2023) reinforces this point, contending that the conditions that make IBE reliable in science—empirical feedback loops, predictive testing, and systematic error correction—are absent in metaphysical contexts. Ritchie (2022) adds that the appeal to theoretical virtues such as simplicity bears no more than a formal analogy between the two domains. If these criticisms are correct, the choice between rival metaphysical explanations risks becoming purely conceptual or intuitive, without the methodological backing that characterizes scientific practice. Whether this constitutes a decisive objection to selective IBE in metaphysics or merely a significant challenge remains an open question in the contemporary debate; what it does suggest, however, is that the selective function of IBE cannot be straightforwardly imported from science to metaphysics without substantial qualification.

The distinction between the selective and the heuristic uses of IBE allows for a more precise understanding of the methodological shift that occurs when the transition is made from science to metaphysics. In the scientific domain, the selective use of IBE relies on criteria such as empirical testability, predictive capacity, and coherence with new observational data. It is a mechanism for choosing among rival hypotheses whose strength depends directly on the possibility of subjecting them to verification and potential refutation. In the metaphysical context, however, these criteria lose their applicability: the absence of empirically testable consequences may prevent the choice between explanations from benefiting from the same corrective process afforded by confrontation with experience—a concern central to the critiques of Saatsi

(2017), Allzén (2023), and Ritchie (2022).

In this scenario, the heuristic use of IBE, closer to what Peirce described as the creative moment of abduction, takes on particular importance. To understand its role, however, it is necessary to specify the mechanism by which abduction operates in the metaphysics of science, rather than merely listing its purported virtues.

The process can be characterized as follows. The metaphysician of science begins with a scientific finding, theory, or structural feature—for instance, the regularity of natural laws, the non-locality of quantum phenomena, or the geometric structure of spacetime. These findings raise ontological questions that the scientific theories themselves describe but do not resolve: What are laws of nature? What underlies quantum non-locality? Is spacetime a substance or a system of relations? Guided by prior philosophical and scientific knowledge, the metaphysician generates a novel ontological hypothesis—a candidate account of the underlying structure that could make sense of the scientific phenomenon in question. This generative step is abductive in the Peircean sense: it introduces a new idea into the explanatory space, it is fallible, and it does not presuppose a prior set of alternatives from which to select.

Consider the metaphysics of laws of nature. Science describes regularities—gravitational attraction, thermodynamic processes, electromagnetic interactions—but the question of what laws of nature ultimately are does not admit of empirical resolution. The various metaphysical proposals that have been advanced—laws as relations between universals (Armstrong 1983), laws as grounded in dispositional properties (Ellis 2009), or laws as ontological primitives—were not selected through a comparative evaluation of pre-existing alternatives guided by empirical feedback. Rather, each was generated as a novel ontological hypothesis in response to features of scientific practice that called for metaphysical articulation. Armstrong's proposal, for example, emerged from the recognition that mere regularities (as defended by regularity theorists) seemed insufficient to account for the modal force and explanatory role that laws play in scientific reasoning; his postulation of laws as second-order relations between universals was an abductive leap aimed at providing an ontological ground for this modal force.

A similar pattern can be identified in the metaphysics of spacetime. General relativity describes the geometric structure of spacetime with remarkable empirical success, but the ontological question of whether spacetime is a substantival entity existing independently of material bodies or a system of spatiotemporal relations between them is not settled by the physical theory itself. Substantivalism and relationalism represent distinct abductive hypotheses, each generated in response to the ontological implications of the physics, and each articulating a different possible structure underlying the empirical content of the theory.

These examples illustrate the mechanism by which abduction operates in the metaphysics of science: it expands the space of ontological possibilities by generating

novel hypothetical frameworks in response to features of our best scientific theories that call for metaphysical elucidation. This is precisely what Morganti & Tahko (2017) characterize as moderately naturalized metaphysics: a form of inquiry that draws on scientific theories and models to explore the realm of metaphysical possibility, unfolding hypothetical scenarios that go beyond the available empirical evidence while remaining informed by and consistent with it. The methodological value of abduction in this context lies not in confirming which ontological framework is true—a task for which, as noted above, the relevant empirical feedback may be unavailable—but in articulating the space of plausible ontological alternatives that our best science leaves open.

Given this characterization of abduction's role, metaphysical explanationism finds its strength not in the selection of "the best" explanation among empirically testable rivals, but in the generation of plausible and fruitful hypotheses that guide conceptual inquiry and expand our understanding of the ontologically possible. This form of explanationism does not offer strong epistemic guarantees of truth—a limitation it shares, to varying degrees, with selective IBE in both scientific and metaphysical contexts—but it does constitute a rational exploratory tool. Importantly, it addresses the concerns raised by Saatsi (2017) and others without requiring a rigid methodological continuity between science and metaphysics: it proposes instead a heuristic continuity, in which abduction serves as the creative bridge between scientific content and metaphysical theorizing.

Our proposal is that IBE in its abductive sense can fulfill a useful constructive role in the metaphysics of science: opening conceptual pathways, articulating plausible scenarios, and delineating the contours of the possible without losing reference to the empirical basis of science. In contrast, applying IBE in its selective sense without empirical mediation faces the significant challenges identified by Saatsi (2017), Allzén (2023), and Ritchie (2022), and its capacity to provide reliable justification in such contexts remains an open question.[6] We believe this clarification lays the groundwork for properly assessing the scope and potential of explanationism in the metaphysics of science, independently of the broader dispute over the epistemic status of selective IBE.

In summary, the status of selective IBE in the metaphysics of science remains a genuinely open question. There are substantial reasons—advanced by Saatsi (2017), Allzén (2023), and Ritchie (2022)—to doubt that its truth-conduciveness can be maintained in the absence of empirical feedback, though defenders of methodological continuity such as Paul (2012) and Williamson (2016) would resist this conclusion. Our proposal does not depend on adjudicating this dispute. What we do claim is that, regardless of the ultimate verdict on selective IBE, understanding abduction as a form of creative, fallible insight provides a clear and defensible account of how metaphysicians of science actually generate their ontological proposals. Abduction,

in this sense, does not overcome the epistemic limitations that affect IBE in metaphysics—it is no less fallible and its products are no less in need of critical scrutiny. But it performs a different function: it operates in the context of discovery rather than the context of justification, and its contribution lies in expanding the space of ontological possibilities rather than in establishing the truth of any particular one. This generative function, we contend, is both clearly operative and clearly valuable in the practice of the metaphysics of science, conceived as an exploration of the realm of possibility (Morganti & Tahko 2017).

5. Conclusion

This paper has argued that the distinction between abduction and selective IBE, often overlooked in the literature, has direct and significant consequences for assessing the role of explanatory reasoning in the metaphysics of science. By disentangling these two senses of IBE—one creative and generative, the other comparative and selective—we have shown that they face different challenges when applied outside the domain of empirical science, and that recognizing this difference opens a productive path for explanationism in metaphysics.

The central philosophical upshot of our argument is the following. The epistemic status of selective IBE in metaphysics is a genuinely open question: there are substantial reasons to doubt that it can retain the truth-conduciveness it enjoys in science (Saatsi 2017; Allzén 2023; Ritchie 2022), though defenders of methodological continuity resist this conclusion (Paul 2012; Williamson 2016). Our proposal does not require settling this broader dispute. What it does show is that, regardless of how the debate over selective IBE is ultimately resolved, abduction in its creative, Peircean sense provides a clearly operative and clearly valuable function in the metaphysics of science. As our analysis of cases such as the metaphysics of laws of nature and the ontology of spacetime has illustrated, metaphysicians of science characteristically generate novel ontological hypotheses in response to features of scientific theories that call for metaphysical elucidation—a process that is abductive in structure, fallible in character, and independent of the comparative selection mechanisms on which selective IBE depends.

This result has implications for how we understand the explanationist program in the metaphysics of science. Explanationism need not be committed to the claim that metaphysical theories can be justified in the same way as scientific theories—through comparative selection guided by empirical feedback. A more modest but defensible form of explanationism grounds the value of metaphysical inquiry in its capacity to expand the space of ontological possibilities, articulating plausible scenarios that are informed by our best scientific knowledge and that enrich our understanding of

what the world could be like at a fundamental level. This aligns with the conception of metaphysics of science as an exploration of the realm of possibility, as defended by Morganti & Tahko (2017), and it is compatible with a range of positions on the broader question of methodological continuity between science and metaphysics.

Several limitations of our proposal should be acknowledged. First, while we have argued that abduction plays a clearly generative role in the metaphysics of science, we have not provided a detailed account of the criteria by which abductively generated hypotheses should be evaluated once they have been proposed. The question of how metaphysicians of science should assess, compare, and refine competing ontological proposals—once the abductive moment has done its work—remains an important open problem that we have not addressed here. Second, our argument has focused on the metaphysics of science specifically, and we do not claim that our conclusions extend without modification to other areas of metaphysics where engagement with scientific theories may be less direct. Third, the agnostic stance we have adopted regarding the epistemic status of selective IBE in metaphysics, while strategically motivated, leaves open substantive questions about whether and how metaphysical claims can ultimately be justified—questions that deserve further investigation.

These limitations suggest several directions for future research. One natural extension of our work would be to investigate whether specific criteria for evaluating abductively generated metaphysical hypotheses can be articulated—criteria that go beyond mere consistency with science but that do not require the full apparatus of empirical testing. Another promising line of inquiry concerns the relationship between the abductive moment in metaphysics and the broader dynamics of scientific change: to what extent do abductively generated metaphysical hypotheses feed back into scientific practice, and can this feedback provide an indirect form of corroboration? Finally, our analysis could be extended to examine other methodological devices—beyond IBE and abduction—that may play analogous roles in bridging scientific and metaphysical inquiry, contributing to a more fine-grained understanding of the methodological landscape of the metaphysics of science.

References

Allzén, S. 2023. "Against Methodological Continuity and Metaphysical Knowledge." *European Journal for Philosophy of Science* 13 (1): Article 5.

Anderson, D. R. 1986. The Evolution of Peirce's Concept of Abduction. In *Transactions of The Charles S. Peirce Society*, 22(2): 145–164.

Armstrong, D. M. 1983. *What is a Law of Nature?* Cambridge University Press. https://doi.org/10.1017/CBO9781139171700

Armstrong, D. M. 1997. *A World of States of Affairs*. Cambridge University Press. https://doi.org/10.1017/CBO9780511583308

Azar, R. M. 2017. ¿Es correcta la identificación entre 'Abducción' e 'Inferencia a la Mejor Explicación'?. *Contrastes. Revista Internacional de Filosofía*, 22(1). https://doi.org/10.24310/Contrastescontrastes.v22i1.3406

Azar, R. M. 2023. Reevaluando el realismo científico: abducción y optimismo epistemológico. *Revista Instante*, 5(3): 216–238. https://doi.org/10.29327/2194248.5.3-11

Azar, R. M. 2019. *La inferencia a la mejor explicación revisitada. Alcances y limitaciones*, Tesis de Doctorado, Universidad de Buenos Aires.

Bryant, A. 2017. "Keep the Chickens Cooped: The Epistemic Inadequacy of Free-Range Metaphysics." *Synthese* 194 (7): 2581–2602.

Campos, D. 2011. On the distinction between Peirce's abduction and Lipton's Inference to the best explanation. *Synthese*, 180(3): 419–442. https://doi.org/10.1007/s11229-009-9709-3

Chakravartty, A. 2017. *Scientific Ontology: Integrating Naturalized Metaphysics and Voluntarist Epistemology*. Oxford: Oxford University Press.

Ellis, B. 2009. *The Metaphysics of Scientific Realism*. Acumen Publishing. https://doi.org/10.4324/9781315711522

Harman, G. H. 1965. The inference to the best explanation. *The philosophical review*, 74(1): 88–95. https://doi.org/10.2307/2183532.

Lipton, P. 2004. *Inference to the Best Explanation*. London: Routledge.

Paul, L. A. 2012. "Metaphysics as Modeling: The Handmaiden's Tale." *Philosophical Studies* 160 (1): 1–29.

Peirce, C. S. 1958. *Collected Papers of Charles Sanders Peirce* (CP), Cambridge, Mass., Harvard University Press. Editado por C. Harstshorne, P. Weiss. y A. Burks.

Peirce, C. S. 1998. *The essential Peirce: Selected philosophical writings 2* (EP2). Indianapolis: Indiana University Press.

Ritchie, J. 2022. "On the Continuity of Metaphysics with Science: Some Scepticism and Some Suggestions." *Metaphilosophy* 53 (2–3): 202–220.

Saatsi, J. 2017. Explanation and Explanationism in Science and Metaphysics. In M. H. Slater; Z. Yudell (Eds.), *Metaphysics and the Philosophy of Science: New Essays* (p.163–191). Oxford University Press.

Sardi, G. C. 2022. Algumas distinções entre a abdução de Charles S. Peirce e a inferência da melhor explicação de Gilbert Harman. *Problemata: International Journal of Philosophy*, 13(2), 62–80. https://doi.org/10.7443/problemata.v13i2.60451

Sardi, G. C. 2024. Inferência da Melhor Explicação, Virtudes Explicativas e Critérios para a Escolha de Teorias. *Principia*, 28(4):639–657. https://doi.org/10.5007/1808-1711.2024.e96386

Sardi, G. C.; Silva, M. R. 2023. A distinção entre geração e aceitação de teorias científicas: um problema para a inferência da melhor explicação. *Griot: Revista de Filosofia*, 23(3): 223–234. https://doi.org/10.31977/grirfi.v23i3.3548

Soto, C. 2014. Una evaluación del realismo científico de Peirce a 100 años de su muerte. *Cuadernos de Filosofía Latinoamericana*, 35(111): 33–58. https://doi.org/10.15332/s0120-8462.2015.0111.01

Soto, C. 2015. The current state of the metaphysics of science debate. *Philosophica*, 90: 23–60. https://doi.org/10.21825/philosophica.82121

Swoyer, C. 2008. Abstract Entities. In T. Sider, J. Hawthorne; D. Zimmerman (Eds.), *Contemporary Debates in Metaphysics* (p.11–31). Blackwell Publishing.

Morganti, M.; Tahko, T. E. (2017). Moderately Naturalistic Metaphysics. *Synthese*, 194(7), 2557–2580. https://doi.org/10.1007/s11229-016-1068-2

Thagard, P. R. 1978. The Best Explanation: Criteria for Theory Choice. *The Journal of Philosophy*, 75(2): 76–92. https://doi.org/10.2307/2025686

Williamson, T. 2016. "Abductive Philosophy." *The Philosophical Forum* 47 (3–4): 263–280.

Notes

[1]Daniel Campos (2011, p.425) also distinguishes two types of abduction within Peirce's work: habitual abduction, in which an already known rule is applied to a new case, and creative abduction, which gives rise to novel hypotheses to explain phenomena hitherto not understood. In both cases, abduction is therefore limited to the so-called context of discovery and not to the context of justification.

[2]In this example (CP 7.498), Peirce is not strictly addressing abduction, but rather the association of similar ideas. Nevertheless, it can be plausibly interpreted as an excellent example of creative abduction.

[3]The central function of IBE is to justify the belief in the truth of a hypothesis based on its superior explanatory power compared to other alternatives. In this sense, IBE belongs to the context of justification, that is, the moment when a hypothesis is evaluated and a decision is made about which hypothesis to adopt regarding a phenomenon.

[4]It is worth noting that the conceptual confusion between IBE and Peirce's abduction originated with Harman himself (1965, p.88). This misunderstanding can be attributed to an inadequate grasp of the development of the concept of abduction within Peircean philosophy. This development is commonly divided into two distinct periods: before and after 1900. In the former, Peirce conceived of the *hypothesis* or *retroduction*—which would later become abduction—as a type of reasoning not clearly distinguished from induction, conceived as a "process of evidencing" (Anderson 1986, p.148) for or against a series of given hypotheses. The conceptual shift occurred around 1903, when he began to treat abduction as an autonomous form of reasoning, exclusively responsible for the creation of hypotheses and preceding any method of experimental verification. His earlier formulations, however, often failed to clearly distinguish between different types of reasoning, partly due to a conception of inference overly guided by the idea of proof—which Peirce himself later recognized as responsible for conflating distinct types of reasoning, as well as for improperly associating abduction both with probabilistic considerations and with syllogistic forms in the study of logic (CP 2.101). It is likely that Harman relied on Peirce's writings prior to 1903, which would explain the conceptual confusion that persisted in the philosophy of science literature, being reproduced even by authors such as Thagard (1978, p.77) and Lipton (2004, p.193), who at times treated the concepts of abduction and IBE as synonymous and at other times failed to adequately distinguish between the processes of theory generation and theory selection.

[5]It is clearly evident that Peirce (CP. 5.197) maintained that all hypotheses generated by abduction can only be considered good hypotheses insofar as they are related to the available observational data and are susceptible to subsequent empirical confirmation within the logic of scientific inquiry, specifically at the stage of induction (that is, at the stage of justification). At first glance, this may suggest that the application of abductions in the meta-

physics of science is not particularly viable, insofar as such hypotheses satisfy the first requirement—namely, being consistent with the available empirical data—but fail to meet the second requirement, that is, being susceptible to subsequent inductive testing (as suggested by IBE). However, it should be emphasized that, according to our position, the application of abduction in philosophy—more specifically in the metaphysics of science—is not subject to the same rules that govern abduction in the scientific context, since it does not fall within the same logic of scientific inquiry in the Peircean sense. In metaphysics in general, empirical confirmation is not achieved in the strict sense; rather, possible explanatory paths are explored. In this context, abduction presents itself as an adequate description of good metaphysical reasoning, functioning as a heuristic guide for the formulation of explanatory hypotheses.

[6]In this sense, we can also say that Lipton's aforementioned first filter of IBE can be partially useful for the metaphysics of science, since, as we have seen, this stage can be interpreted as a type of abduction that allows the initial identification or generation of plausible hypotheses based on background knowledge and prior expectations, without immediately requiring a rigorous comparison among rivals or empirical confirmation. However, we cannot affirm that the entire structure of Lipton's IBE serves the purposes of the metaphysics of science, since the second filter in this model is, indeed, a stage of deliberate comparison among competing hypotheses, which involves a systematic evaluation of their explanatory virtues. This characterizes a selective process that goes beyond simple creative generation and requires empirical confirmations—thus being similar to Harman's IBE. Therefore, it is fundamental to distinguish these two stages to understand the limits and potentials of IBE and abduction in the context of a moderately naturalized metaphysics of science.