Weak and Local Methodological Continuity across Science, Naturalized, and Autonomous Metaphysics

Bruno Borge
University of Buenos Aires, CONICET

Forthcoming in *Synthese* (special issue: *The Aims and Methodologies of Contemporary Metaphysics*). Accepted April 2026. Please cite the published version when available.

Abstract: This paper defends weak and local methodological continuity—that is, continuity with respect to specific shared devices rather than global methodological frameworks—across science, naturalized metaphysics, and autonomous metaphysics. The contemporary debate has conflated two distinct continuity problems: whether naturalized metaphysics can share in science's epistemic success and whether proximity to science privileges certain metaphysical approaches over others. Additionally, continuity claims involve multiple dimensions that are not always properly distinguished. The paper makes two interconnected contributions. First, I develop a triadic framework that distinguishes these continuity problems and dimensions, providing analytical tools for practice-sensitive assessments. Second, applying this framework, I argue that thought experiments, understood as shared methodological devices, exhibit methodological continuity across all three domains. The analysis shows that thought experiments perform structurally analogous roles in each domain despite operating under different epistemic constraints and background commitments. This methodological continuity neither presupposes nor entails strong epistemic conclusions, but demonstrates that the science-metaphysics relationship resists simple dichotomies of complete continuity or categorical discontinuity. The framework suggests a productive methodological reorientation: rather than debating global continuity a priori, philosophical inquiry should examine how specific methodological devices function across domains through careful, local case studies. Thought experiments provide one such case; the approach developed here applies to other shared methodological tools.

Keywords: metaphysics, methodological continuity, thought experiments, naturalized metaphysics, science and metaphysics, epistemic warrant

1. Introduction

The relationship between science and metaphysics has become a central concern in contemporary analytic philosophy, especially within the philosophy of science. At stake is whether metaphysical inquiry can legitimately claim continuity with scientific practice, and if so, what form that continuity should take. The question has generated sharply divergent answers. On one end, radical naturalizers such as Ladyman and Ross (2007) argue that metaphysics is epistemically legitimate only insofar as it is effectively absorbed into science. On the other, defenders of autonomous or "armchair" metaphysics maintain that metaphysical theorizing retains a distinctively a priori methodology independent of empirical constraints.

A middle position is articulated by Morganti and Tahko (2017), who defend what they call *moderately naturalistic metaphysics*. On their view, science and metaphysics share their subject matter—both aim to describe the structure of reality—but remain methodologically distinct. Metaphysics retains an irreducible a priori methodology oriented toward the exploration of possibility space and the articulation of highly general concepts such as essence, dependence, and modality. While metaphysical inquiry must be constrained by science and plays an indispensable role in interpreting scientific theories, its methods are not continuous with scientific methods in any strong sense. Metaphysical hypotheses are formulated, assessed, and compared using distinctive criteria that cannot be reduced to empirical confirmation, prediction, or experimental control. Morganti and Tahko thus reject both radical naturalization and armchair isolationism. Other influential voices, however, defend methodological continuity. Paul (2012) conceives of metaphysics as model-building continuous with scientific practice, while Williamson (2016) argues for a broadly abductive methodology shared across philosophy and science. These positions, though differing in scope and emphasis, converge on the claim that metaphysical and scientific theorizing proceed by structurally similar methods.

While global claims about methodological continuity or discontinuity between science and metaphysics remain open to dispute, a more fine-grained assessment is both possible and necessary—one that is sensitive to the diversity of both scientific and metaphysical practice. The present paper defends the view that methodological continuity, understood in a weak and local sense, can be established by focusing on specific methodological devices rather than comprehensive methodological strategies. Making this case, however, requires distinctions more refined than those that have characterized what I will call *the continuity problem* in the literature thus far. By identifying methodological continuity at a local level, I aim to show that such continuity can obtain across all three poles—science, naturalized metaphysics, and autonomous metaphysics.

The argument proceeds as follows. Section 2 surveys the contemporary debate, diagnosing a persistent conflation: what has been treated as a single continuity problem is better understood as involving multiple distinct questions that have not been systematically distinguished. I identify two clusters of positions, each organized around different concerns—one addressing methodological affinity between science and metaphysics broadly, the other addressing epistemic delimitation within metaphysics. While these clusters do not map cleanly onto a single partition, they reveal systematic patterns that motivate a more fine-grained framework. Section 3 responds to this diagnosis by developing a triadic conceptual framework that distinguishes three poles of inquiry (science, naturalized metaphysics, autonomous metaphysics). This framework reveals two distinct continuity problems: the relationship between science and naturalized metaphysics, and the relationship between naturalized and autonomous metaphysics. The framework also explores three dimensions along which continuity may be assessed (methodological, ontological, epistemic). Together, these distinctions provide analytical tools for clarifying the questions raised—but often conflated—by both clusters. Section 4 applies this framework by defending the core claim: thought experiments, understood as shared methodological devices, exhibit weak and local methodological continuity across all three poles. This continuity is compatible with significant discontinuities elsewhere, and it neither presupposes nor entails strong epistemic conclusions. The upshot is a more nuanced picture of the science-metaphysics relationship—one that resists both sweeping continuity claims and categorical discontinuity theses.

Together, these sections advance two interconnected contributions: a refined analytical framework for continuity debates, and a substantive claim about methodological continuity that this framework makes visible. Crucially, these contributions are not merely juxtaposed but logically complementary. The framework reveals that continuity questions are best pursued locally and dimension by dimension; the case study demonstrates that this local strategy yields a determinate, non-trivial result. Conversely, the case study retrospectively vindicates the framework by showing that the distinctions it introduces—between poles, dimensions, and levels of analysis—are not idle taxonomic machinery but tools that do genuine philosophical work.

2. The science-metaphysics continuity problem

2.1. The big picture

Metaphysics, even restricted to the analytic tradition, encompasses many traditions, problems and methodologies. Traditionally, what is often called *armchair, standard* or *autonomous metaphysics* has been conceived as a philosophical inquiry conducted primarily on the basis of a priori considerations. On this conception, metaphysics proceeds through methodological devices such as conceptual analysis, intuitions, modal reasoning, and a priori argumentation. This idea is deeply associated with the very practice of metaphysics, and that is no coincidence; this is the main way metaphysics has been practiced for centuries and is a popular trend in ongoing debates.

Naturalized metaphysics, *metaphysics of science*, or *scientific metaphysics*, among others, are general labels that encompass a series of attempts to relate classic approaches to metaphysical inquiry (conceived as centrally a priori) to scientific inquiry (conceived as centrally a posteriori). Many of these approaches relate to the idea of a naturalization of metaphysics, in which the solution to some or all metaphysical problems is deferred to science. Prominently, Ladyman and Ross (2007) argue for a complete naturalization, since "standard analytic metaphysics . . . contributes nothing to human knowledge" (2007, p. 1). Others are guided by the far less radical intuition that investigations into some or all metaphysical problems should be informed by the outcomes of science. The sense in which science should inform metaphysics can be cashed out in many ways, but all of them impose some kind of restrictions on speculative theorizing stemming from scientific knowledge. The spectrum encompasses several possibilities that range from strong dependence on (or even reduction to) science to forms of collaborative coexistence.

Debates concerning how, and to what extent, metaphysics can or should be naturalized or scientifically informed have been articulated in a variety of ways. Yet they are overwhelmingly framed in terms of the question of continuity between science and metaphysics. This framing presupposes that both science and metaphysics are epistemic enterprises aimed at the production of legitimate, factual knowledge, and it is precisely this shared epistemic aspiration that motivates the continuity question. The central issue, then, is whether science and metaphysics should be regarded as continuous or independent disciplines and, if some form of continuity is to be defended, how that continuity is to be characterized.

Within this general continuity framework, Morganti and Tahko (2017) propose a useful distinction between two dimensions along which the continuity between science and metaphysics can be assessed: whether they share their methods and whether they share their subject matter—or, as I will call it here, their ontological domain. Combining these dimensions yields several possible views concerning the science–metaphysics continuity. Morganti and Tahko's *moderately naturalistic metaphysics* considers that science and metaphysics are continuous at the level of subject matter, insofar as both aim to describe the structure of reality, while remaining methodologically distinct, insofar as the methods of metaphysics are essentially a priori.

However, many authors are reluctant to admit the methodological discontinuity pointed out by Morganti and Tahko. On such views, metaphysical theorizing proceeds in ways closely analogous to scientific theorizing, often by means of abductive reasoning guided by theoretical virtues. Paul (2012), while denying ontological continuity, defends methodological continuity: she conceives of metaphysics as model-building continuous with scientific practice. On Paul's account, metaphysical theories—like scientific theories—are developed by constructing models (involving idealization, abstraction, and hypothetical systems) which are then compared and selected using inference to the best explanation (IBE). The theoretical virtues employed in this selection process—simplicity, ontological parsimony, elegance, explanatory power, and fertility—are the same virtues deployed in scientific theory choice (Paul 2012, p. 12). While the categories of entities represented in metaphysical models differ from those in scientific models (properties and substances rather than particles or genes), the methodological strategy of model construction, comparison, and abductive selection remains structurally continuous across both domains. Similarly, Williamson (2016) defends a broadly abductive methodology shared across philosophy and science. In a more radical vein, Ladyman and Ross (2007) argue that metaphysics is legitimate only insofar as it is guided by the methods of science.

Morganti and Tahko are not alone in their skepticism about methodological continuity between science and metaphysics. However, the debate over continuity takes different forms depending on what is primarily at stake. While Morganti and Tahko's critique is embedded in a broader discussion about how metaphysics and science should best relate—defending metaphysical autonomy while maintaining ontological continuity—another prominent line of criticism focuses more directly on the epistemic legitimacy of autonomous metaphysics in comparison to naturalized approaches. Critics such as Allzén (2023), Bryant (2017), and Ritchie (2022) argue that the apparent methodological similarities between science and metaphysics—most notably the appeal to IBE and/or to theoretical virtues in theory selection—do not suffice to ground an epistemically robust continuity between the two domains. On their view, the justificatory role these methods play in science crucially depends on their integration with empirical testing and feedback, a feature that metaphysical inquiry lacks.

At the same time, other authors have sought to articulate more moderate or hybrid positions, according to which metaphysics can be scientifically informed without collapsing into science, or can fruitfully interact with science while retaining a degree of methodological autonomy. Views of this kind are developed, in different ways, by Chakravartty (2017), Lee (2023), and Emery (2025), who respectively emphasize empirical constraint, a division-of-labour model, and a stronger normative link between scientific content and scientific methodology. While Lee (2023), as noted above, denies strong ontological continuity, his division of labour model nonetheless allows for fruitful interaction between analytic and scientific metaphysics.

Taken together, these debates suggest that what is often presented as a single dispute could be better understood as involving multiple, partially independent questions. These include not only whether science and metaphysics are methodologically continuous, but also how different forms of metaphysical inquiry relate to one another, and whether the very notion of "continuity" requires further disambiguation.

2.2. Two clusters of problems in the contemporary debate

Disputes around the continuity problem are born of a primarily epistemic concern. It is assumed that science is a successful epistemic enterprise, and that this success is deeply related to its methodological means. Questions about metaphysics arise naturally in this context: if both science and metaphysics are in the business of finding truths about the world, what is the relationship between them? The debate typically focuses on whether they share the same object of inquiry, whether they employ similar methodologies, and whether methodological or ontological continuity suffices to secure epistemic success for metaphysics. A further question concerns the scope of such continuity within metaphysics itself. If continuity with science is what grounds epistemic legitimacy, this suggests a potential divide between those forms of metaphysical inquiry that are closely informed by science and those that proceed more autonomously. The status of this latter, more autonomous metaphysics then becomes unclear: is it simply epistemically deficient, or does it play a different but nonetheless legitimate role?

The debate engages these questions, but does so in ways that are not systematically integrated. Considered as a whole, the literature is dispersed around specific points of interest rather than organized around a unified framework. However, there are two clusters of arguments and discussions that have developed, each organized around different concerns. One addresses methodological affinity between science and metaphysics broadly, the other addressing which forms of metaphysical inquiry—if any—can legitimately claim epistemic warrant, and does proximity to scientific practice provide a principled basis for such delimitation.

The first cluster addresses methodological affinity between science and metaphysics broadly construed. Authors in this cluster examine whether metaphysical inquiry as such employs methods continuous with those of science. The central questions concern whether science and metaphysics share similar strategies of theory construction and selection, and whether these methodological similarities suffice to establish genuine continuity.

Some authors in this cluster defend methodological continuity between science and metaphysics, though they differ in scope and emphasis. Paul (2012), as noted in section 2, conceives of metaphysics as model-building continuous with scientific practice, arguing that both domains employ similar strategies of theory construction and evaluation guided by theoretical virtues. Williamson (2016) extends this claim even further. He is concerned not only with metaphysics but with philosophy as a whole, including logic. This connects with anti-exceptionalist projects, which regard logic as epistemically and methodologically continuous with empirical science. On this view, there is nothing special about logic that exempts it from the methods of science; rather, logical theories are selected by means of

IBE, just as scientific theories are. Logic, metaphysics, and science thus form a methodological continuum, unified by shared inferential practices.

Brenner (2023) offers a further defense of methodological continuity along related lines. Using a Bayesian framework, he argues that appeals to theoretical virtues in metaphysical theory choice are vindicated by their structural analogue in science, and that objections alleging a relevant disanalogy between the two domains rest on a misdescription of how theoretical virtues actually function in scientific practice. Like Paul and Williamson, however, Brenner treats metaphysics as a unified domain without distinguishing between its more and less scientifically informed variants. A more internally differentiated picture emerges from Schurz (2020), who defends a program of "inductive metaphysics" in which metaphysical theories are justified by creative abduction. Crucially, Schurz proposes two rationality conditions that distinguish scientific from speculative abductions: achievement of unification and independent testability. This internal distinction within abductive methodology exerts pressure toward a finer-grained framework: Schurz's own conditions imply that not all metaphysical abductions are on a par, and that the epistemic credentials of metaphysical inference depend on whether it satisfies constraints that are characteristically scientific. His framework thus anticipates, from within this cluster, the need for the kind of triadic differentiation developed in Section 3.

On the skeptical side, several influential voices challenge the claim of methodological continuity. Morganti and Tahko (2017), as noted in Section 2, defend methodological discontinuity between science and metaphysics, arguing that metaphysics retains an irreducible a priori dimension that prevents genuine methodological continuity with science. Allzén (2023) develops a more targeted critique of the reasoning underlying Paul's and Williamson's defenses of continuity. The main target of his criticism is the inference from the reliability of abductive methodology in science to its reliability in metaphysics: even if both domains employ similar explanatory reasoning, this does not establish that IBE plays the same role in each case. Allzén maintains that the uses of IBE in science and the uses of IBE in metaphysics are distinct with only superficial similarities, since the conditions that make IBE reliable in science—empirical feedback loops, predictive testing, and systematic error correction—are absent in metaphysical contexts.

Ritchie (2022) develops a related critique with respect to the use of IBE in both fields and the appeal to theoretical virtues such as simplicity: “These forms of reasoning have no more than a formal analogue in metaphysics. It is false to claim that if simplicity judgements are truth conducive in biology or other sciences, they are truth conducive in metaphysics” (2022: 208). Once again, although issues of epistemic warrant are central to this line of argument, the strategy neither relies on nor aims at establishing an intrinsic delimitation within metaphysics on the basis of epistemic legitimacy.

What unifies this cluster is a focus on methodological comparison at a general level, treating metaphysics as a unified domain rather than distinguishing systematically between its more and less scientifically informed variants. However, a closer inspection reveals a significant ambiguity. When Paul argues that metaphysics proceeds by model-building and abductive selection, or when Williamson defends a broadly abductive methodology shared across philosophy and science, the paradigmatic examples they invoke—possible worlds reasoning, mereological theories, debates about universals and tropes—are drawn almost exclusively from autonomous metaphysics. Their claims about methodological continuity thus target, in practice, a specific form of metaphysical inquiry, even as they are formulated in general terms. This matters because the question of whether *autonomous* metaphysics is methodologically continuous with science is a different question from whether *naturalized* metaphysics—which is by design responsive to scientific practice—is so continuous. The first question is contentious precisely because autonomous metaphysics does not obviously share science's empirical constraints; the second is contentious for the opposite reason, namely that naturalized metaphysics may be *too* close to science to constitute a genuinely distinct enterprise. By treating metaphysics as a unified domain, this cluster conflates questions that may well have different answers.

The second cluster addresses epistemic delimitation within metaphysics. Authors in this cluster do distinguish—at least implicitly—between forms of metaphysics more or less closely tied to scientific practice, and ask which of these forms can legitimately claim epistemic credentials. To address this question, authors in this cluster typically examine the relationship between science and naturalized metaphysics as a means of assessing whether and to what extent proximity to science confers epistemic warrant. The positions range from eliminativist to moderately permissive, but share a common concern with internal demarcation within metaphysics.

Eliminativist positions, such as Ladyman & Ross (2007), argue that only metaphysics reducible to or continuous with physics is epistemically legitimate; autonomous metaphysics should be eliminated. A skeptical position is defended by Melnyk (2013), who explicitly acknowledges the possibility that even naturalized metaphysics may fail to constitute a viable form of inquiry. While holding that science—particularly fundamental physics—is the most promising (and perhaps the only promising) resource for addressing metaphysical questions, Melnyk does not assume that this suffices to secure the epistemic legitimacy of metaphysics as such. His view is therefore best understood as a comparative claim: if any metaphysical approach is to succeed, it must be continuous with science, yet it remains an open question whether even such an approach ultimately succeeds. Melnyk's position is particularly instructive for the present purposes. Unlike most authors in the first cluster, who treat "metaphysics" as a unified domain and ask whether it is continuous with science, Melnyk explicitly confronts the possibility that even within naturalized metaphysics, the relationship with science is not straightforward. His skepticism thus presupposes precisely the kind of internal differentiation within metaphysics that the first cluster leaves unexamined.

A more moderate stance is articulated by Bryant (2017). Bryant argues that what she calls "free range metaphysics" is epistemically inadequate insofar as it operates almost exclusively with criteria such as consistency, simplicity, intuition, and explanatory power, which are not truth-conducive. Nevertheless, she maintains that free-range metaphysics should not be eliminated, since it has important collateral benefits, such as providing clarity, contributing to conceptual incubation, and supplying formal tools that can enrich philosophical reflection. The resulting normative conclusion partly aligns with that of Ladyman and Ross, but is significantly more moderate. On Bryant's view, the problem is not that free-range metaphysics exists, but rather that it is taken in bad faith as a producer of justified metaphysical knowledge.

More moderate positions, such as Chakravartty's (2017), turn the continuity problem into a matter of degree, understood in terms of epistemic risk. In Chakravartty's framework, there is a continuum regarding epistemic risk that goes from science to pure metaphysics; every point in that continuum implies a different degree of epistemic risk. As we move from statements that are more empirically vulnerable to others that are less empirically vulnerable, we are increasing epistemic risk. Within metaphysics, there is no given or privileged way of distinguishing where to draw the line between the metaphysics that is sufficiently warranted by experience and that which is not. What counts as too much epistemic risk can reasonably vary across different frameworks.

By contrast, Humphreys (2013) appeals to a general risk-averse epistemic principle, according to which inquiry should aim to minimize epistemic risk whenever possible. On this basis, he maintains that ontological commitments grounded in scientific practice are epistemically preferable to those arising from more speculative forms of metaphysical theorizing. The claim, however, is comparatively modest: scientific ontology—or naturalized metaphysics—is said to involve less epistemic risk than autonomous or speculative metaphysics, without offering a precise account of how different ontological approaches are ordered along a risk continuum, nor of where the threshold between acceptable and unacceptable epistemic risk should be drawn.

Jaksland (2023a) exposes these internal tensions systematically. He identifies a trilemma for radical naturalized metaphysics: either it provides a wholesale argument for its epistemic credibility—risking reliance on the very traditional methods it criticizes—or it makes the weaker claim that science-based metaphysics is merely superior to other metaphysics—which may not suffice to establish its own epistemic legitimacy—or it replaces its eliminative stance with a more permissive view that allows autonomous metaphysics wherever science has no bearing, thus undermining its own eliminative ambition. Jaksland (2023b) further argues that the underdetermination of metaphysics by science forces naturalized metaphysics into an increasingly restrictive position in which metaphysicians are effectively displaced rather than rescued. These internal difficulties reinforce the observation that the relationship between science and metaphysics cannot be adequately assessed through a single continuity question, and they motivate the kind of fine-grained analysis developed in Section 3.

This cluster is unified by concerns about epistemic delimitation: determining which parts of metaphysics, if any, possess epistemic legitimacy, typically by examining whether naturalized metaphysics can inherit epistemic credentials from science and whether autonomous metaphysics is thereby epistemically deficient. While methodological considerations arise, they are subordinate to the primary epistemic concern.

These two clusters, while addressing different concerns, share a common limitation: neither attends systematically to the distinct comparisons their arguments presuppose. The first cluster treats metaphysics as a unified domain when asking about methodological continuity with science, yet the epistemic significance of any such continuity plausibly depends on what kind of metaphysics we are considering—whether it is already informed by science or proceeds more autonomously. The second cluster does distinguish between more and less scientifically informed forms of metaphysics when assessing epistemic legitimacy, and it productively examines how proximity to science might confer epistemic warrant. Yet the relationship between science and naturalized metaphysics is often assessed in relatively coarse terms, without fully articulating the dimensions along which continuity is claimed—whether methodological, ontological, or epistemic—or the conditions under which such continuity would obtain. As a result, while this cluster advances important questions about epistemic delimitation, it leaves open how exactly the relevant forms of continuity should be characterized.

Moreover, both clusters tend to conflate continuity along different dimensions. Questions about methodological overlap, ontological overlap, and epistemic parity are often run together without explicit distinction. The result is that disagreements appearing to concern 'the continuity problem' often involve participants addressing different questions or operating with different understandings of what continuity amounts to.

What is needed, then, is a more fine-grained conceptual framework that carefully distinguishes not only the poles of comparison but also the dimensions along which continuity is claimed, and the distinct continuity problems these comparisons generate. The next section develops such a framework. As I will argue in section 4, this framework not only clarifies the existing debate but also opens a path for making progress on it: once the relevant distinctions are in place, it becomes possible to assess continuity locally—at the level of specific methodological devices—rather than as an all-or-nothing verdict on the science-metaphysics relationship.

3. Towards a more fine-grained framework

3.1 Distinguishing two continuity problems

The survey in the previous section identified two clusters of positions, each addressing different aspects of the science-metaphysics relationship, not always properly distinguishing the dimensions along which continuity might obtain. A key source of this confusion, I suggest, is the failure to distinguish between three poles of comparison: science, naturalized metaphysics, and autonomous metaphysics. This triadic distinction is motivated by a structural observation: the two clusters examined above are, in effect, addressing different relata. The first cluster asks whether metaphysics is methodologically continuous with science, but its central arguments implicitly concern autonomous metaphysical practices. The second cluster examines whether proximity to science confers epistemic legitimacy, which presupposes a distinction between metaphysics that is scientifically informed and metaphysics that is not. Making these three poles explicit reveals that what has been treated as a single continuity problem is better understood as involving (at least) two logically independent problems.

The first problem concerns the relationship between science and naturalized metaphysics: Can metaphysical inquiry that takes science as its primary point of reference legitimately share in science's epistemic success? This problem asks whether naturalized metaphysics—understood as a form of inquiry explicitly constrained by and responsive to scientific practice—can inherit epistemic warrant from science, and if so, under what conditions and to what extent. Importantly, this is not a question about metaphysics as such, but about a specific way of doing metaphysics.

The second continuity problem concerns the relationship between naturalized and autonomous metaphysics: Does proximity to science provide a principled basis for making epistemologically relevant distinctions between different forms of metaphysical inquiry? This is an intra-metaphysical problem that asks whether naturalized metaphysics should be regarded as epistemically or methodologically privileged over more autonomous, a priori approaches. It raises questions about the normative weight of scientific constraint within metaphysics: What evaluative consequences follow from adopting (or rejecting) scientific constraints? What roles, if any, can autonomous metaphysics legitimately play?

Crucially, these two continuity problems are logically independent. One may consistently hold that naturalized metaphysics is continuous with science while denying that autonomous metaphysics stands in any relevant continuity relation with naturalized metaphysics. This can be exemplified by Ladyman and Ross' eliminativism. Conversely, one may defend forms of interaction or continuity between naturalized and autonomous metaphysics while remaining skeptical about whether naturalized metaphysics genuinely shares in science's epistemic credentials. While this combination is less commonly defended, Lee's (2023) division-of-labour model approximates this structure: analytic metaphysics can provide abstract models useful for scientific metaphysics (affirming interaction between the two), yet this does not require that naturalized metaphysics directly inherits epistemic warrant from science in the way defenders of strong continuity claim. Finally, one may reject continuity along both axes, or affirm it along both. The logical independence of these two problems is further corroborated by recent work exposing internal tensions within radical naturalized positions: Jaksland's (2023a) trilemma shows that the relationship between science and naturalized metaphysics cannot be straightforwardly leveraged to settle the status of autonomous metaphysics, since each horn of the trilemma leaves the second continuity problem unresolved. Recognizing this independence helps to clarify why agreement on one continuity claim does not automatically extend to the other, and why debates on continuity often generate cross-cutting and seemingly unstable alignments.

Before turning to the dimensions along which continuity may be assessed, a further distinction is worth noting. Continuity questions can be formulated at two distinct levels of analysis: descriptive and normative. At a descriptive level, one can ask what types of continuity science and metaphysics actually exhibit, and discover that, despite appearances, scientists and metaphysicians work in significantly similar ways in some respect—or, on the contrary, that beyond superficial similarities they are engaged in radically different enterprises. At a normative level, one can ask whether that is the way things *should* be done. The answers may range from confirming that current practices should remain unchanged, to suggesting that some of them (e.g., a priori metaphysical reasoning) should be discontinued or performed without expectation of epistemic gains. The answers at the descriptive level are an important input to questions at the normative level. Current debates often conflate, or at least do not fully distinguish, both levels when reflecting on continuity issues. The weak and local methodological continuity defended in section 4 is intended as a descriptive claim: a characterization of how the three poles actually operate with respect to one specific methodological device. The corresponding normative question remains open; section 4.3 returns briefly to this issue.

3.2 Dimensions of continuity: methodological, ontological, and epistemic

The triadic framework and the distinction between two continuity problems enable finer-grained analysis than has typically been available. But specifying the relata of continuity claims is only part of the story, since there is more than one sense in which continuity can obtain. The central point that this section develops is the following: different dimensions of continuity may obtain to different degrees across different connections in the triad. Science and naturalized metaphysics may be continuous along one dimension—say, ontological—while being discontinuous along another—say, methodological. And the pattern of continuity and discontinuity between science and naturalized metaphysics need not mirror that between naturalized and autonomous metaphysics. This means that a complete assessment of the science-metaphysics relationship requires specifying not only the poles being compared but also the dimension along which continuity is claimed. The dimensions most prominently discussed in the literature, building on Morganti and Tahko's (2017) influential distinction, are continuity with respect to the object of inquiry and continuity with respect to methods. But a third dimension—epistemic continuity—has been implicitly at work throughout the debate and deserves explicit treatment.[1]

Continuity with respect to object of inquiry—or ontological continuity—was central to the positions in the literature surveyed above: Paul and Lee, for instance, deny ontological continuity by distinguishing between categories (investigated by metaphysics) and instances of categories (investigated by science), while Morganti and Tahko defend it by arguing that both aim to describe the structure of reality.

Discussions on methodological continuity are prolific and have been extensively illustrated in the previous sections; however, the relationship between the methodological and the strictly epistemic dimensions of the continuity problem is rarely unpacked, and both are frequently conflated. A careful distinction can be helpful.

Two disciplines may share identical or highly similar methods while nonetheless differing significantly with respect to the degree of epistemic reliability or justification those methods afford. The mere fact that the same method is employed across different domains does not guarantee that it will play the same epistemic role in each case. This point can become clearer if epistemic continuity itself is disambiguated. In a broad sense, epistemic continuity may be understood simply as the claim that two disciplines are both capable of producing epistemically warranted knowledge. In a stricter sense, however, epistemic continuity involves more than this: it requires that the epistemic warrant produced in each domain be secured in a sufficiently similar way, typically through the successful use of shared methods. Without this stricter requirement, epistemic continuity would be trivially satisfied by any two epistemically successful enterprises—meteorology and constitutional law, for instance, would count as epistemically continuous merely in virtue of both producing warranted knowledge, despite sharing neither methods nor subject matter. It is this stricter sense that is

[1] One advantage of this multi-dimensional approach is that it helps characterize the three poles themselves—and especially naturalized metaphysics, whose extension often suffers from ambiguity—not through stipulative definitions but through their relational profiles. On this view, naturalized metaphysics is distinguished from science by retaining a degree of methodological and ontological autonomy (it addresses distinctively metaphysical questions, even if constrained by scientific findings), and from autonomous metaphysics by its responsiveness to empirical evidence and its acceptance of science as a substantive constraint on theorizing. This relational characterization avoids the need for a sharp demarcation while providing enough structure for the assessment of continuity claims.

most often implicitly assumed in continuity arguments, and it is precisely here that methodological and epistemic continuity risk being conflated.

Consider IBE. Even if both science and metaphysics employ IBE as a methodological device, this does not entail that IBE is equally truth-conducive or epistemically reliable in both domains. That is precisely the complaint raised by Allzén (2023) against Paul's continuity claims: the reasons typically offered by scientific realists for trusting IBE in science may be absent or significantly weakened in metaphysical contexts. This shows that methodological continuity is in principle distinct from epistemic continuity, understood as the production of comparable epistemic warrant.

It is worth noting that epistemic continuity differs from the other two dimensions in an important respect: it has an irreducibly evaluative character. Ontological and methodological continuity are, in the first instance, descriptive claims—they assert that two domains share their subject matter or their methods, without thereby evaluating whether either domain is epistemically successful. Epistemic continuity, by contrast, functions as what might be called a *success term*: to say that two domains are epistemically continuous is to say not merely that they operate similarly but that they achieve comparable epistemic results. This asymmetry reinforces the importance of distinguishing these dimensions rather than conflating them. It also helps explain why methodological continuity alone does not entail epistemic continuity: the former describes a structural similarity, while the latter registers an epistemic achievement that depends on further conditions—such as the availability of empirical feedback—that may obtain in one domain but not in another.

The three dimensions distinguished above—ontological, methodological, and epistemic—are not exhaustive. For specific analytical purposes, it may be fruitful to identify additional dimensions, such as pragmatic continuity (concerning shared investigative practices rather than shared methodological principles) or normative continuity (concerning transfers of normative force between domains).[2] While the analysis in Section 4 focuses exclusively on the methodological dimension, the framework is in principle open to extension along these and other dimensions.

The framework developed in this section raises a question it is uniquely positioned to address. If continuity between science and metaphysics admits of multiple dimensions, and if global claims along any single dimension remain deeply disputed, a natural next step is to ask whether continuity can be established locally—that is, with respect to specific methodological devices rather than comprehensive methodological strategies. The triadic structure is essential to this move: it allows us to test whether a given device operates analogously across all three poles, rather than collapsing the inquiry into a binary science-versus-metaphysics comparison that obscures the intermediate position of naturalized metaphysics. Section 4 pursues this strategy. It first develops non-triviality conditions for methodological continuity claims (section 4.1), and then argues that thought

[2] Pragmatic continuity would concern whether different domains engage in shared investigative *practices*—such as styles of reasoning in Hacking's sense or what Suárez (2024) calls a "modeling attitude"—rather than sharing methodological frameworks or devices. Normative continuity would concern whether there is a transfer of normative force from one domain to another; for instance, if metaphysics identifies ultimate categories that constrain which entities science may investigate (see Varga 2021 for a related discussion). Neither dimension collapses into ontological, methodological, or epistemic continuity.

experiments, understood as shared methodological devices, exhibit weak and local methodological continuity across all three poles (section 4.2). Thought experiments are a particularly apt test case for this purpose: unlike abductive theory choice—where disputes about continuity turn on the presence or absence of empirical feedback mechanisms—they display structurally analogous roles across domains while remaining comparatively less sensitive to the domain-specific epistemic constraints that make global continuity claims contentious.

4. Weak and local methodological continuity

4.1 Non-triviality requirements for methodological continuity

Before defending the claim that there is a sense in which methodological continuity obtains between science, naturalized metaphysics, and autonomous metaphysics, it is necessary to clarify what should count as a philosophically substantive continuity claim. It is worth flagging at the outset that the continuity defended here operates at the level of specific methodological devices rather than comprehensive methodological frameworks; the non-triviality conditions developed below are designed to make this restriction precise. Not every similarity in epistemic practice suffices to establish methodological continuity in a non-trivial sense. In particular, continuity should not be understood as identity of methods across the board, nor as a wholesale transfer of justificatory force from one domain to another. Two distinct disciplines may employ similar forms of reasoning or evaluation while differing significantly in their aims, background assumptions, and epistemic constraints.

This point is especially important given that many epistemic practices, insofar as they pursue epistemic aims at all, will inevitably rely on general forms of reasoning, such as inference, evaluation of evidence, or comparative assessment of hypotheses. Under such a permissive understanding, virtually any two epistemic practices—for instance, ornithology and social epistemology—could be said to be methodologically continuous. While such claims may be indisputable, they are also philosophically uninformative. If methodological continuity is to play any substantive role in the analysis of the relationship between science and metaphysics, it must therefore satisfy some sort of non-triviality requirement.

In what follows, I take *substantive methodological continuity* to require more than the mere presence of generic epistemic practices. In order to avoid triviality, a claim of methodological continuity must satisfy at least one of two conditions. First, two disciplines may exhibit a broad affinity in their methodological orientation—concerning their general approach to inquiry, the kinds of resources they treat as evidentially relevant, and the styles of reasoning they employ—in a way that sets them apart from other disciplines within their respective fields. Second, two disciplines may share specific methodological devices or tools that play an analogous role within their respective practices—that is, devices that are employed in structurally similar ways, serve comparable functions, and aim at similar epistemic or explanatory goals within their respective contexts of inquiry.

Two clarifications about the notion of method at work here are in order. First, it is important to distinguish between methods understood as *comprehensive methodological frameworks*—encompassing epistemic norms, justificatory standards, evidential constraints, and reliability conditions—and methods understood as *specific methodological devices*—concrete tools, techniques, or patterns of reasoning employed within inquiry. The continuity debate in the literature has predominantly operated at the level of comprehensive frameworks: Paul's modeling strategy, Williamson's abductive methodology, and the skeptical responses of Allzén and Ritchie all concern whether science and metaphysics share a general methodological orientation. The two non-triviality conditions proposed here cut across this distinction. Methodological affinity (the first condition) concerns similarities at the level of general orientation—it asks whether two disciplines share an approach to inquiry that sets them apart from other fields. Shared devices (the second condition) operates at a different level: it asks whether specific tools play analogous roles across disciplines, regardless of whether those disciplines share a broader methodological orientation. This distinction matters because the continuity I will defend is of the second kind—local, device-level continuity—and it can obtain even between disciplines that differ substantially in their overarching epistemic norms and justificatory structures.

Second, a shared device, as I understand it here, involves more than the bare fact that two disciplines employ a tool that can be described under a common label. That would be too permissive: both physics and literary criticism "use arguments," but this shared description is too thin to establish methodological continuity. A device is genuinely shared, in the relevant sense, when it satisfies three conditions: (i) it is employed in structurally similar ways—that is, the sequence of operations it involves (setup, execution, assessment) follows a recognizably common pattern; (ii) it serves comparable functions within the respective practices—for instance, testing theoretical commitments, exploring consequences, or reorganizing explanatory priorities; and (iii) its outputs are subject to analogous modes of criticism—that is, practitioners in each domain can challenge the device's deployment by the same kinds of moves (revising premises, questioning idealizations, proposing counter-scenarios). When a device satisfies these three conditions across domains, the resulting continuity is substantive rather than nominal, even if the epistemic constraints governing each domain differ significantly.

Pure cases—where one condition is fully satisfied while the other is entirely absent—are rare in general, but especially in scientific practice. Most disciplinary relationships exhibit some degree of both methodological affinity and shared devices. Nevertheless, examining cases where the balance differs significantly helps clarify the distinction and illustrates how each condition can contribute independently to substantive methodological continuity.

Consider the relationship between climatology and macroeconomics. Both disciplines investigate highly complex systems characterized by multiple interacting components, nonlinear dynamics, and emergent properties. They rely heavily on idealized models—often computational—and operate under conditions of significant uncertainty and limited empirical access. Neither discipline can intervene experimentally on the system as a whole: climatologists cannot rerun Earth's atmosphere under controlled conditions, and macroeconomists cannot experimentally manipulate national economies. In both cases, explanatory and predictive success is mediated by simulations, scenario analysis, and statistical inference rather than by controlled experimentation or direct causal intervention. Both fields have developed characteristic practices of projecting future system states conditional on stipulated boundary conditions—emission scenarios in climate science, growth and policy assumptions in macroeconomics—and both face the challenge that their models cannot be straightforwardly validated against the very futures they project. This yields a clear affinity in methodological orientation, one that distinguishes both disciplines from more experimentally driven sciences such as molecular biology or materials physics. While climatology and macroeconomics do employ some structurally similar tools—both use dynamic models with feedback loops and equilibrium assumptions—these shared devices are relatively generic features of complex systems modeling rather than specific methodological innovations transferred between the fields. The distinctive methodological affinity lies not in particular shared tools but in a broader convergence of approach, inferential style, and evidential resources. In this case, the first condition predominates over the second[3].

An example with the opposite profile can be found in the use of optimality models across behavioral ecology and rational choice theory in economics. In both domains, such models explore the consequences of assuming that agents optimize a given quantity under constraints—fitness in ecology, utility in economics. The mathematical structure is closely analogous: both derive predictions about behavior from optimization principles, employ similar formal techniques (e.g., Lagrangian methods, dynamic programming), and serve comparable functions within their respective theoretical frameworks. In both cases, the optimization operates in an "as if" sense: organisms do not consciously maximize fitness, just as rational choice models need not presuppose that agents consciously compute utilities—yet the formal apparatus is the same. The methodological device—optimality modeling—is genuinely shared and plays structurally analogous roles. Nevertheless, the broader methodological orientation of the disciplines differs substantially. Behavioral ecology embeds these models within evolutionary explanation, tests them against observational data from natural populations, and treats deviations from optimality as puzzles requiring adaptive or phylogenetic explanation. Rational choice theory in economics, by contrast, often treats optimality models as part of normative theory with implications for decision-making, relies on different forms of evidence (experimental, survey, market data), and addresses optimization failures through behavioral modifications of the core framework. The underlying causal mechanism also diverges: natural selection in one case, rational deliberation (or its bounded approximations) in the other. The empirical grounding, explanatory goals, and disciplinary context diverge significantly even as the formal tools overlap. Here, the second condition predominates over the first[4].

Taken together, these examples illustrate two important points. First, the proposed non-triviality conditions genuinely discriminate between different kinds of methodological overlap: not every shared epistemic feature qualifies as substantive continuity. Second, they support the claim that methodological continuity can be stated in different senses. Disciplines may be methodologically continuous in one respect but not in another. This

---

[3] The epistemological challenges distinctive of simulation-based science are discussed in Winsberg (2018); for a treatment of model-based reasoning and scenario analysis as central inferential strategies in economics, see Morgan (2012, esp. chs. 6–7). On the interplay between climate modeling and economic projections, and the shared reliance on computational models for systems that resist controlled experimentation, see Gramelsberger and Feichter (2011).

[4] On the structural parallels between fitness-maximization in evolutionary biology and utility-maximization in rational choice theory, and on the conditions under which these parallels are and are not sustained, see Okasha (2018, esp. secs. 6.6–6.8 and ch. 8). For the broader topic of model transfer between economics and biology, see Weisberg, Okasha, and Mäki (2011) and the accompanying essays in that special issue.

reinforces the idea that interesting continuity claims must be carefully specified and assessed case by case, rather than asserted at the level of whole disciplines or methods.

The two conditions are independent: neither is necessary nor sufficient on its own, and disciplines may satisfy one without satisfying the other. Substantive methodological continuity admits of different strengths. *Strong continuity* would require satisfying both conditions—exhibiting broad methodological affinity while also sharing specific devices. *Weak continuity*, by contrast, requires satisfying only one of the two conditions.

The continuity defended in this paper is weak in precisely this sense. Additionally, it is *local*, in the sense that it just satisfies the second condition. This conception of continuity is deliberately modest. It neither presupposes strong epistemic conclusions nor excludes the possibility of significant discontinuities elsewhere. As argued in section 3.2, methodological and epistemic continuity are logically distinct dimensions. The claim defended here concerns only the methodological dimension.

A potential objection at this point is that the notion of "shared device" captures continuity in argumentative form or dialectical function rather than in method in any epistemically robust sense. After all, if methodology in philosophy of science is typically understood as involving justificatory standards, evidential constraints, and reliability conditions, then showing that thought experiments perform structurally analogous roles across domains may fall short of establishing methodological continuity in the sense that matters. This objection has force against continuity claims formulated at the level of comprehensive methodological frameworks—and it is precisely the objection that Allzén and Ritchie raise against Paul's and Williamson's defenses of continuity through IBE. However, the continuity defended here operates at a different level. It concerns specific devices, not comprehensive strategies. The claim is not that science and metaphysics share a general method of inquiry, but that they share a particular tool whose deployment follows a recognizably common pattern and whose outputs are subject to analogous modes of criticism. This is a thinner notion of methodology than what is at stake in the IBE debate, but it is not thereby trivial: as the non-triviality conditions above are designed to show, not every shared practice qualifies. The deliberate modesty of the claim is what insulates it from the objections that correctly target stronger continuity theses.

Whether there is global methodological continuity across the three poles remains an ongoing debate, as the previous sections intend to show. Besides its complexities, one of the main reasons for not addressing the point here is that any diagnosis on global continuity depends crucially on how naturalized metaphysics is conceived. On radical naturalistic conceptions, global continuity with science seems to be easier to secure, but at the risk of collapsing metaphysics into science, thereby eliminating any substantive distinction between the two. As Ladyman himself admits, radical naturalization is compatible with the view that "naturalized metaphysics is nothing more than addressing traditional metaphysical questions . . . in a way that is continuous with theoretical debates in science (though one may want to say that this is just science, not metaphysics—that is a merely terminological issue)" (Ladyman 2017, 149). On more moderate conceptions, where naturalized metaphysics is more obviously distinct from science, methodological continuity remains deeply disputed.

Rather than trying to settle these disagreements here, I adopt a different strategy: examining continuity locally by analyzing specific methodological devices that are neutral with respect to how naturalized metaphysics is conceived. The remainder of this section illustrates this approach through a detailed analysis of thought experiments, specifically aiming to demonstrate how this particular device satisfies the second non-triviality condition by operating analogously across different domains.

4.2 Thought experiments as a test case

The remainder of this section defends the claim that thought experiments constitute a shared methodological device across science, naturalized metaphysics, and autonomous metaphysics, thereby establishing a weak and local form of methodological continuity. This claim bears on both continuity problems distinguished in section 3.1—the relation between science and naturalized metaphysics, and the relation between naturalized and autonomous metaphysics—in the same way: in each comparison, there is continuity with respect to this specific device, satisfying the second non-triviality condition but not necessarily the first. The case study proceeds in three steps: first, establishing that thought experiments are indeed deployed across all three domains; second, analyzing their functional role within each domain; and third, showing that this role is structurally analogous despite differences.

Thought experiments are used across all three domains of our triadic framework: autonomous metaphysics, naturalized metaphysics, and science. This is simply a descriptive fact about contemporary practice. In autonomous metaphysics, thought experiments such as Putnam's Twin Earth or Parfit's fission cases are central to theory evaluation. In naturalized metaphysics, philosophers deploy thought experiments to explore questions about laws, causation, persistence, and composition in dialogue with scientific theories. In science, thought experiments like Galileo's falling bodies or Einstein's elevator famously played crucial roles in theoretical development. While I will center the defense of my main thesis on the analysis of thought experiments, I believe that this general strategy for exploring local methodological continuity can be applied to the analysis of other ubiquitous methodological devices (see section 4.3 for discussion).

For the purposes of this paper, a minimal working characterization of thought experiments will suffice. A thought experiment is a narrated experimental set-up designed to be carried out in imagination rather than materially. Its scenario specifies (often idealized) initial conditions and operations, invites the reader to "run" the set-up mentally, and aims at answering a definite theoretical question by tracking what would happen (or what would follow) under those suppositions. Thought experiments can be realizable in principle, but their epistemic point is not exhausted by realizability; they are often deployed precisely when real experimentation is unavailable or impractical. One of the central philosophical questions around thought experiments is how (if at all) they can support epistemically significant conclusions, especially in the absence of new empirical input. In the contemporary literature, this question has generated a well-known family of disputes, including (at least) the contrast between "argument views", according to which thought experiments ultimately function as arguments that can be made explicit, and more robust accounts on which some thought experiments deliver a distinctive kind of insight not reducible to ordinary argumentation (cf. Brown 1991; Norton 1991, 1996; Sorensen 1992; Gendler 2000; Frappier, Meynell & Brown 2013).

The adoption of this minimal, practice-based characterization calls for explicit justification, since one might worry that a sufficiently thin notion of thought experiment builds cross-domain portability into the definition rather than deriving it from substantive analysis. Three points address this concern. First, the characterization is minimal but not vacuous: it specifies structural features—a narrated scenario, idealized conditions, a mental operation, and a theoretical conclusion—that are jointly sufficient to distinguish thought experiments from other argumentative or heuristic devices (such as analogies, classification schemes, or purely formal proofs), while remaining neutral on disputed questions about their epistemic status. Second, the continuity claim defended below does not rest on the characterization alone but on a detailed functional analysis (Section 4.2) showing that thought experiments satisfy the non-triviality conditions specified in Section 4.1. A thin characterization that merely established that the same label applies across domains would not suffice; what matters is that the devices so characterized are deployed in structurally similar ways and serve comparable functions. Third, the thesis is robust across competing accounts of what thought experiments are. On Norton's argument view, thought experiments are disguised arguments whose force derives from their logical structure; if so, the structural parallels documented in Section 4.2 reflect genuine argumentative parallels rather than superficial narrative resemblance. On Brown's rationalist view, some thought experiments provide non-empirical access to abstract features of reality; if so, the cross-domain recurrence of thought experimentation is not merely a sociological fact but reflects a shared capacity that operates independently of domain-specific empirical constraints—which, if anything, strengthens the continuity claim. The minimal characterization adopted here is thus not a way of avoiding these debates but a way of identifying a common explanandum that all parties recognize, while leaving the deeper explanatory questions open.

Even amid disagreement about their epistemic status, there is a useful cluster of points that enjoy broad acceptance and that can be productive for the local-continuity strategy pursued here. First, and perhaps centrally, the function of thought experiments as a methodological device is generally contended as highly similar (if not identical) in science and philosophy, and the default attitude in the specialized debates is “that there is no categorical difference between scientific and philosophical thought experiments” (Brown & Fehige 2023). With differences, many authors present unified accounts that reject domain-based essentialism about thought experimentation and emphasize the portability of thought experiments across scientific and philosophical practice, focusing on their not necessarily empirical character (Brown 2004, Cooper 2005); their adaptability to different types of background commitments (Goffi & Roux 2018, p. 451) and other methodological and epistemological reasons (Atkinson and Peijnenburg 2003, pp. 317–18; Boniolo 1997; Gähde 2000; Sorensen 1992, pp. 11–15. See Wilkes 1988 for an exception). Although scientific and philosophical thought experiments are distinguished primarily by their context of application, sometimes this is not enough: some thought experiments that are used in scientific contexts are highly metaphysically loaded, while others that appear in philosophical contexts strongly rely on scientific outputs (cf. Cooper 2005, p. 329). For instance, quantum measurement scenarios straddle physics and metaphysics, while Twin Earth—paradigmatically philosophical—depends on chemical theory. Second, thought experiments characteristically proceed by subjunctive (often causal) idealization: they ask what would happen if certain features were held fixed, varied, or bracketed, and they rely on systematic distortions and abstractions to isolate the dependencies of interest. Third, thought experiments are not automatically successful: their results can be undermined by underspecification, illicit idealization, or the importation of assumptions that the target theory does not license. These are familiar, practice-sensitive failure modes, and they already suggest why the epistemic standing of thought experiments is not something that can be settled globally and a priori.

Several influential taxonomies classify thought experiments by their evidential or dialectical role. Brown’s (1991) well-known division distinguishes *destructive* from *constructive* uses, while Sorensen (1992) offers a complementary classification in terms of refutation format. Although these schemes are not equivalent, they overlap in highlighting the centrality of negative/refutational uses. Importantly, the point of invoking this sort of classificatory apparatus is not to privilege any single taxonomy, nor to take a stand on the ultimate epistemic credentials of thought experimentation. It is simply to introduce a thin comparative lens: if both destructive and constructive deployments recur across science, naturalized metaphysics, and autonomous metaphysics, then thought experiments provide an especially clear case in which weak, local methodological continuity can be made visible.

In the sciences, destructive thought experiments often function as streamlined refutations of entrenched theoretical packages by showing that they generate implausible consequences when pushed in controlled, imaginary settings. Galileo’s falling bodies case is a canonical illustration: by considering tied objects (or equivalently, a compound body) under the Aristotelian assumption that heavier bodies fall faster, the scenario generates an unstable pair of predictions—one implying that the composite should fall faster (because it is heavier), and another implying that it should fall slower (because the lighter body would “drag” the heavier one). The point is not that the case is observationally decisive on its own, but that it enacts a recognizably destructive methodological role: it exposes a structural inconsistency in a candidate principle when combined with minimal auxiliary assumptions, thereby motivating rejection or revision. Einstein's elevator thought experiment illustrates the constructive role in science. By imagining an observer in a uniformly accelerating elevator, the scenario makes salient the equivalence between gravitational and inertial effects—an equivalence that classical frameworks treated as merely accidental. The thought experiment does not refute a specific theory but reorganizes the conceptual landscape: it motivates a shift from treating gravity as a force to understanding it as a manifestation of spacetime curvature. Here the constructive function is clear: the thought experiment guides positive theorizing by making explicit a structural dependency that existing frameworks obscured.

In naturalized metaphysics, thought experiments frequently function at the interface between metaphysical claims and scientifically informed constraints. A prominent destructive use is found in Kim’s causal exclusion case. The core setup combines (i) the causal completeness of the physical, (ii) the supervenience of the mental on the physical, and (iii) a ban on systematic overdetermination. Within that structured space, non-reductive mental causation appears to be squeezed out: if every physical effect already has a sufficient physical cause, then positing an additional irreducible mental cause threatens to produce either redundancy (overdetermination) or causal impotence. Here the method is recognizably destructive in the relevant sense: by constructing an idealized dependence structure and showing that rival options cannot jointly satisfy a set of commitments treated as highly plausible (and in part empirically motivated), the thought experiment functions as a diagnostic refutation of certain metaphysical packages.[5] As Ritchie observes, the exclusion model "allows us to explore the

[5] Kim's exclusion case is discussed as a thought experiment in the literature (see, e.g., Ritchie 2022, p. 217). While it is more schematic than paradigmatic cases involving concrete imagined scenarios, its

possible connections between various weakly theorized concepts, like causation and supervenience" (2022, p. 217). Constructive thought experiments in naturalized metaphysics guide positive theorizing by suggesting metaphysical requirements for accommodating scientific practice. Tooley's (1977, p. 669) particle interaction scenario illustrates this: imagine ten types of fundamental particles where 54 of 55 possible two-particle interactions have laws, but particles X and Y, due to their distribution, can never interact. Is there nonetheless a law governing X-Y interactions? A regularity theory would deny this (the generalization is vacuously true), yet it seems reasonable to affirm it. The thought experiment thus motivates a robust conception of laws of nature. As Brown (1994, pp. 97–98) observes, Tooley's scenario connects directly with the structure of the standard model in particle physics: the model exhibits second-order regularities—patterns among patterns across particle families—that suggest there could be laws governing ever heavier quarks even if such laws are never instantiated. The constructive function is clear: by making salient the pressure to posit uninstantiated laws, the scenario guides metaphysical theory toward accounts that can accommodate nomological realism within a scientifically informed framework.

In autonomous metaphysics, destructive thought experiments are routinely used to undermine analyses by producing counterexamples that exploit modal or conceptual commitments taken to be stable within the practice. Putnam’s Twin Earth case is a standard example. By holding fixed the internal states of speakers while varying the external environment in which a term is used, the thought experiment presses against internalist accounts of meaning (and, by extension, certain pictures of reference and content) by aiming to show that sameness “in the head” does not guarantee sameness of extension or content. Again, the methodological point is not that the case is self-validating, but that it implements the same destructive move as in the scientific examples: it isolates a theoretically controlled scenario that forces a choice among commitments and thereby motivates rejection or revision of a candidate view. Constructive thought experiments in autonomous metaphysics often play a complementary role: they motivate positive reconceptualizations by making salient a structure that existing frameworks obscure. Parfit’s fission/splitting cases can be framed in this way. By exploring scenarios in which psychological continuity and connectedness come apart from strict identity, the cases aim to shift theoretical focus from identity to the relations that matter for survival, responsibility, or prudential concern. Even if one disputes the evidential force of the conclusion, the methodological function is clear: the thought experiment operates as a constructive model that reorganizes explanatory priorities and clarifies what a theory must capture. Briefly, one can also mention more purely clarificatory cases (e.g., Ship of Theseus–style puzzles) as additional constructive tools; but for present purposes, the Twin Earth/Parfit pairing already suffices to exhibit the same destructive/constructive pattern found in science and in naturalized metaphysics.

Across these examples, a structural pattern emerges. Destructive thought experiments—whether Galileo's falling bodies, Kim's exclusion model, or Putnam's Twin Earth—share a common methodological structure. Each constructs an idealized scenario that combines a target principle with background commitments treated as stable, then shows that the combination generates an inconsistency, contradiction, or intolerable consequence. The methodological role is analogous across contexts: to isolate a theoretical vulnerability by

---

methodological function—constructing an idealized dependence structure to derive consequences that motivate theoretical revision—is characteristic of thought experimentation.

controlling for confounding factors and making the problematic dependency explicit. While the background commitments differ, the structural operation remains constant: first, construct a controlled scenario; second, combine it with commitments treated as stable; third, derive an unacceptable result, thereby motivating revision or rejection. Constructive cases exhibit a parallel structure. Einstein's elevator, Tooley's particle interaction scenario, and Parfit's fission cases all reorganize the theoretical landscape by making salient dependencies or relations that existing frameworks obscure, thereby guiding positive theory construction. The function is comparable: not to refute directly but to reconfigure possibility space in ways that render certain principles compelling or certain explanatory priorities clear. That these methodological moves recur across domains with different epistemic constraints—empirical testing in science, scientific grounding in naturalized metaphysics, conceptual coherence in autonomous metaphysics—demonstrates that the device itself is portable and context-adaptable rather than domain-specific. This is precisely what the second non-triviality condition requires: shared devices employed in structurally similar ways to serve comparable functions within their respective epistemic contexts.

This is the limited—but philosophically instructive—payoff of the case study: it displays a shared device with an analogous functional profile across domains, thereby making weak, local methodological continuity visible without settling any global continuity thesis. More generally, it illustrates how the triadic framework can be used to ask continuity questions in a practice-sensitive way, by focusing on specific tools rather than whole disciplines.

A note on the choice of examples is in order. The thought experiments analyzed above—Galileo's falling bodies, Einstein's elevator, Kim's exclusion argument, Tooley's particle scenario, Putnam's Twin Earth, Parfit's fission cases—are all highly canonical. One might worry that this introduces a selection bias: these are precisely the cases that philosophers have repeatedly invoked to make science and philosophy look structurally similar, and they may not be representative of ordinary practice, particularly in contemporary science where thought experiments are often embedded in heavy formal and institutional infrastructures. This concern can be addressed on two fronts. First, canonical cases are canonical precisely *because* they exhibit the structural features of thought experimentation with particular clarity—the idealized scenario, the controlled operation, the derivation of a theoretically significant result. They function in the philosophical literature much as paradigm cases function in Kuhnian analysis: not as statistically representative samples but as instances that make the relevant features maximally visible. The claim is not that all thought experiments across all domains look exactly like these, but that these cases display, in a perspicuous form, the structural pattern whose cross-domain recurrence is the basis of the continuity claim. Second, the analysis already includes diversity along several dimensions that matter for the argument: the examples span three distinct domains, include both destructive and constructive uses, and range from cases with heavy empirical content (Galileo, Einstein) to cases with minimal empirical dependence (Parfit). What they share is the structural pattern identified in Section 4.2—and it is the recurrence of that pattern, not the fame of the examples, that grounds the continuity thesis.

4.3 Addressing objections

Having defended the claim that thought experiments exhibit weak and local methodological continuity across science, naturalized metaphysics, and autonomous metaphysics, I will address three potential objections that might undermine or limit the scope of this conclusion.

A first natural objection is that thought experiments in science are empirically constrained in ways that thought experiments in metaphysics are not, making them fundamentally different devices. In other words, arguing for local continuity through thought experiments might be no different from arguing for global continuity by relying on abductive methods. As previously mentioned, Allzén has plausibly argued that realist defenses of IBE in science depend crucially on "empirical considerations, predictive power and inductive evidence, all of which are paradigmatically absent in the metaphysical context. Without the kind of empirical success that includes 'the generation of novel predictions that are in principle testable' (Psillos 1999), there is no way to justify abductive theory selection" (2023, p. 17-18). If thought experiments as employed in autonomous or naturalized metaphysics lack the empirical constraints present in their scientific uses, the similarities among them might be superficial and the claims of local methodological continuity based on those similarities can be challenged.

However, there are several reasons why this is not a pressing worry for the kind of continuity claims defended in this work. As shown in section 4.2, there is an established consensus in the literature that thought experiments in science and philosophy are not categorically different, functioning similarly despite differences in context and background commitments. Additionally, a closer comparison reveals a structural disanalogy between the two cases that explains why the objection against IBE does not transfer to thought experiments. In abductive theory choice, the method itself involves selecting among competing theories on the basis of explanatory virtues. The reliability of this selection procedure depends on conditions external to the method—empirical feedback loops, predictive success, systematic error correction—that operate *downstream*, on the products of the inference. When these conditions are absent, as Allzén and Ritchie argue they are in metaphysical contexts, the method loses its claim to reliability even if its formal structure remains the same. Thought experiments, by contrast, do not select among competing theories on the basis of explanatory virtues. They explore the consequences of specified background commitments within a controlled imagined scenario. Their evidential force depends not on downstream empirical feedback but on the cogency of the setup, the stability of the background commitments, and the validity of the inference from scenario to conclusion. The key point, then, concerns where empirical constraints operate. Even if thought experiments are not empirically constrained in the sense of testability, other empirical constraints apply. In thought experiments, the empirical constraint is upstream, in the premises from which the thought experiment departs, not downstream in the product of the thought experiment or in its capacity to generate novel testable predictions. Crucially, these kinds of constraints are not present only in scientific thought experiments while being absent in philosophical contexts. Thought experiments in physics, naturalized metaphysics, and autonomous metaphysics depart from background assumptions and explicit setups that are to different degrees empirically (or even scientifically) informed. Consider Putnam's Twin Earth case, a paradigmatic instance of autonomous metaphysics. The thought experiment depends on the assumption that a substance can be superficially identical to water while differing in molecular structure (H_2O vs. XYZ). While the specific chemical details are not essential—any substance with a hidden compositional difference would serve the same argumentative purpose—the scenario nonetheless presupposes a broadly scientific picture of the world in which superficial properties can come apart from underlying structure. The empirical content is not incidental decoration but part of what makes the thought experiment's conclusion compelling. Yet it generates no novel testable predictions about chemistry or reference. Even in cases that clearly belong to the domain of autonomous metaphysics, then, background assumptions can carry substantial empirical/scientific content. This variation in the type and degree of empirical constraint on the premises is compatible with continuity of the methodological device precisely because that device is not in the business of generating the kind of novel testable predictions that Allzén correctly identifies as absent in abductive methodology.

A related concern is that, even if thought experiments function analogously across domains, they may carry significantly different *weight* within each practice. In autonomous metaphysics, where direct empirical testing is unavailable, thought experiments arguably play a more central evidential role than in science, where they typically serve as auxiliaries to empirical investigation. The EPR thought experiment, for instance, was eventually overtaken by Bell's theorem and subsequent experimental tests—something that has no parallel for Parfit's fission cases. This asymmetry is real, but it does not undermine the continuity claim defended here. What is at stake is not whether thought experiments have the same epistemic significance in every domain—that would be a claim about epistemic continuity, which the paper explicitly brackets—but whether they are deployed as the same kind of methodological device: one that constructs controlled scenarios, combines them with background commitments, and derives results assessable through analogous modes of criticism. A hammer plays a more central role in carpentry than in geology, but it remains the same tool employed in a recognizably similar way. The differential centrality of thought experiments across domains is a fact about disciplinary structure, not about the identity of the device.

A second objection concedes the distinction between methodological and epistemic continuity defended in section 3.2, but questions its philosophical significance. Even if these dimensions can be adequately distinguished, one might worry that establishing merely methodological continuity is philosophically uninteresting if it does not "pay dividends" at the epistemic level. This worry is compounded by the fact that debates within the philosophy of thought experiments have not reached consensus on whether thought experiments possess any substantive epistemic power—and the present paper is not the place to resolve that question. Nevertheless, two points are worth considering. First, as emphasized in section 3.2, methodological continuity—even when isolated from epistemic questions—can be philosophically relevant to understanding the nature of a discipline, particularly in contexts where three distinct forms of inquiry are being articulated and distinguished. Second, without settling the question of the epistemic credentials of thought experiments, there are positive indications that the case of thought experiments may exhibit not only methodological but also a form of weak epistemic continuity. Two considerations support this second point. On the one hand, in each of the three domains, a substantial body of literature defends the epistemic credentials of thought experiments. While there are also critics, the persistence of sustained philosophical defense across multiple theoretical frameworks and traditions suggests that thought experiments are not obviously epistemically inert. If they were, the debate would not have the structure it does. On the other hand, an examination of concrete historical cases shows that the formulation of thought experiments had tangible impact on theory adoption and rejection across all three domains. In science, Galileo's and Einstein's thought experiments influenced theoretical development. In metaphysics (both autonomous and naturalized), thought experiments have similarly shaped which views are taken seriously and which are abandoned. While this does not prove that thought experiments have genuine epistemic value (people can be systematically wrong), it does show that epistemic communities ascribe epistemic value to this device. The fact that practitioners across domains treat thought experiments as epistemically relevant is itself a fact that underpins the plausibility of a weak epistemic continuity. It is worth recalling the distinction drawn in section 3.1 between descriptive and normative levels of analysis. The claim defended here operates at the descriptive level: thought experiments *are* employed as shared methodological devices across domains. The normative question—whether practitioners *should* employ them, and under what constraints—remains open, and its answer may well differ across domains. Establishing methodological continuity at the descriptive level does not prejudge this normative issue, but it does provide the factual basis any such normative assessment would require.

A sharper version of this worry holds that the weak and local continuity thesis is not merely philosophically modest but dialectically inert—that is, compatible with every major position in the debate, including those of strong skeptics such as Allzén and Ritchie. On this reading, even authors who deny methodological continuity between science and metaphysics would accept that practitioners in both domains use thought experiments, and that these experiments look structurally similar. If so, the thesis establishes nothing that any party to the debate would contest.

This objection, however, rests on a conflation between acknowledging a superficial similarity and accepting the substantive claim defended here. The thesis is not merely that thought experiments *occur* in both domains—that is indeed uncontroversial—but that they satisfy specific non-triviality conditions: they are deployed in structurally similar ways, serve comparable functions, and are subject to analogous modes of criticism. This is a claim with determinate content that not all positions can easily accommodate. Consider three consequences.

First, the result is incompatible with strong discontinuity theses that characterize science and metaphysics as sharing *no* genuine methodological resources. If thought experiments are genuinely shared devices in the sense specified by the non-triviality conditions, then the science-metaphysics relationship cannot be one of categorical methodological discontinuity. Positions such as Morganti and Tahko's, which maintain that the methods of metaphysics are "essentially a priori" and not continuous with scientific methods, must at minimum be qualified: whatever discontinuity obtains at the level of comprehensive methodological frameworks, it does not extend to all specific devices.

Second, the result constrains skeptical positions. Allzén and Ritchie correctly argue that IBE-based continuity claims fail because the conditions that make IBE reliable in science are absent in metaphysics. But this argument depends on a specific feature of IBE—its reliance on downstream empirical feedback. It does not generalize to all methodological devices. The present case study shows that thought experiments, precisely because their evidential force depends on upstream commitments rather than downstream feedback, are not vulnerable to this objection. A skeptic who wishes to maintain that there is no philosophically significant methodological overlap between science and metaphysics must now address this case specifically, rather than relying on the IBE argument as a general-purpose objection to continuity.

Schurz's (2020) framework for inductive metaphysics helps sharpen this point. Schurz proposes that rational abduction requires both unification and independent testability—conditions that the skeptics are right to find absent in much autonomous metaphysical practice. But these conditions apply specifically to abductive inference, where the selection among competing theories depends on downstream empirical feedback. Thought experiments, as argued above, do not operate by abductive selection; their evidential force depends on upstream commitments rather than downstream testability. The weak and local continuity defended here thus identifies what survives at the methodological-dispositional level once stronger epistemic conditions such as Schurz's are acknowledged: a form of continuity that obtains precisely where the standard skeptical arguments—designed for abductive methodology—lose their grip.

Third, and most importantly, the result reframes the terms of the debate. If weak and local methodological continuity can be established for at least one device, then the question shifts from "Is there continuity between science and metaphysics?" to "For which devices does continuity obtain, and what are the conditions under which it does?" This is a more tractable and empirically informed question than the global one that has generated stalemates in the literature. The framework and case study together suggest that the productive path forward is not to argue a priori about continuity in general, but to build a map of local continuities and discontinuities across the triadic structure.

A third and final objection targets the scope and generalizability of the argument. A natural worry is that, even if thought experiments do exemplify weak, local methodological continuity, they remain just one case—perhaps a peculiarity that says more about thought experimentation than about the science-metaphysics relation. In the vein of the point made in the previous paragraph, this concern can be addressed by noting that thought experiments are plausibly one instance of a broader family of cross-domain methodological devices. Similar continuity claims can be tested by examining, for example, (i) model construction and idealization techniques (such as toy models or targetless models) as controlled distortions that isolate theoretical dependencies, (ii) robustness reasoning via systematic variation of assumptions to assess the stability of conclusions, (iii) symmetry and invariance considerations as constraints on admissible representations and ontological commitments, and (iv) formal methods—including axiomatization, proof techniques, and semantic modeling—as tools for making commitments explicit and exploring their consequences. The present case study does not license sweeping generalizations from a single example. Rather, it illustrates a more general strategy: examine how specific methodological devices operate under different background commitments and epistemic constraints, and assess continuity at that local, practice-sensitive level.

5. Conclusion

This paper has defended a modest but substantive claim: thought experiments exhibit weak and local methodological continuity across science, naturalized metaphysics, and autonomous metaphysics. The case for this claim rests on a distinctive feature of thought experimentation that sets it apart from other contested methodologies. Unlike abductive theory choice—where disputes about methodological continuity turn on the presence or absence of empirical feedback mechanisms, predictive testing, and systematic error correction—thought experiments display structurally analogous methodological roles across domains. They construct a highly idealized scenario, invite the reader to perform a constrained operation within it, and extract a result that can be criticized by revising the set-up, challenging the inference, or proposing counter-thought experiments. What matters epistemically is not the acquisition of new empirical data but the controlled unfolding of consequences from background commitments—empirical, theoretical, or conceptual—already in place.

This explains why objections targeting methodological continuity arguments based on abductive methodology do not transfer straightforwardly to thought experimentation. Even in paradigmatic scientific cases, thought experiments are often not directly testable as described and their effectiveness does not depend on empirical predictions or experimental testability. Assessment proceeds instead via coherence with accepted theory and the robustness of outcomes under variations of the imagined set-up. Thought experiments thus occupy a distinctive methodological position: they depend on background assumptions that may include empirical content, but their characteristic evidential force is comparatively less sensitive to domain-specific empirical constraints than abductive theory choice. This makes them especially revealing for assessing methodological continuity.

The detailed analysis in section 4 shows that thought experiments, understood as specific methodological devices, are genuinely shared tools across domains—not merely superficially similar practices. The structural pattern is constant: destructive cases isolate theoretical vulnerabilities by deriving unacceptable consequences from target principles combined with stable commitments; constructive cases reorganize theoretical landscapes by making salient dependencies or relations that guide positive theorizing. This vindicates the triadic framework developed in section 3 and the distinction between methodological and epistemic dimensions: we can identify methodological overlap without thereby settling epistemic questions about whether these shared methods yield comparable epistemic warrant across domains. Importantly, this is not a claim about global methodological continuity, nor does it resolve foundational epistemic debates about thought experiments or metaphysics more broadly.

The modest payoff, however, is philosophically significant in at least three respects. First, the result is incompatible with strong discontinuity theses: if thought experiments are genuinely shared devices satisfying the non-triviality conditions specified in Section 4.1, then the science-metaphysics relationship cannot be one of categorical methodological discontinuity, even if comprehensive methodological frameworks diverge substantially. Second, the result constrains how skeptical arguments can be formulated. Objections to methodological continuity based on IBE—such as those developed by Allzén and Ritchie—depend on features specific to abductive theory choice (downstream empirical feedback, predictive testing) that do not generalize to all methodological devices. The case of thought experiments shows that some devices are insulated from these objections, which means that the skeptic must engage with continuity claims device by device rather than relying on a single general-purpose argument. Third, and most importantly, the result suggests a productive reorientation of the debate. If weak and local continuity can be established for at least one device, the question shifts from whether science and metaphysics are continuous *tout court* to which devices exhibit continuity, under what conditions, and with what consequences. This is a more tractable and empirically informed question than the global one that has generated stalemates in the literature. Rather than arguing a priori about continuity in general, philosophical inquiry should build a map of local continuities and discontinuities across the triadic structure. Thought experiments provide the first entry on that map. The framework developed here invites others.

The triadic framework developed in this paper has further applications beyond the case study pursued here. By making explicit the three poles and the two continuity problems, it provides diagnostic tools for analyzing tensions that existing positions leave unresolved. For instance, moderate naturalism as defended by Morganti and Tahko maintains that science and metaphysics share their subject matter but not their methods. The framework makes visible a question this position does not address: if methodological discontinuity holds at the level of comprehensive frameworks, does it also hold at the level of every specific device? The present case study suggests it does not—and this creates an internal tension for moderate naturalism that merits further investigation. Similarly, the framework can sharpen eliminativist positions: Ladyman and Ross's claim that only metaphysics continuous with science is epistemically legitimate gains precision when reformulated in triadic terms, since the framework reveals that continuity may obtain along some dimensions and with respect to some devices while failing along others. Indeed, the internal tensions exposed by Jaksland (2023a, 2023b) within radical naturalized positions, and the contrast between Schurz's (2020) demanding epistemic conditions and the device-level continuity established here, already illustrate the diagnostic work that the triadic framework can perform beyond the scope of the present case study. Whether this undermines or refines the eliminativist program is a question the framework makes tractable but that exceeds the scope of this paper. These are among the paths that the triadic structure opens. The present contribution has pursued only one of them—establishing a local continuity result for a single device. But if the framework proves useful for these broader purposes, the modesty of the initial result will have been not a limitation but a deliberate starting point.

Acknowledgements. An earlier version of this paper was presented at the Philosophy Department Colloquium, University of Miami. I am grateful to Otávio Bueno, Anjan Chakravartty, Aleksandra Hernandez, and the audience for valuable comments and discussion. I also thank two anonymous reviewers for *Synthese*, whose careful and constructive feedback substantially improved the paper.

References

Allzén, S. (2023). Against methodological continuity and metaphysical knowledge. *European Journal for Philosophy of Science, 13*(1), Article 5. https://doi.org/10.1007/s13194-022-00505-6

Atkinson, D., & Peijnenburg, J. (2003). When are thought experiments poor ones? *Journal for General Philosophy of Science, 34*(2), 305–322.

Boniolo, G. (1997). On a unified theory of models and thought experiments in natural sciences. *International Studies in the Philosophy of Science, 11*(2), 121–142. https://doi.org/10.1080/02698599708573558

Brenner, A. (2023). Theoretical virtues and the methodological analogy between science and metaphysics. *Synthese*, 201, Article 54. https://doi.org/10.1007/s11229-023-04047-z

Brown, J. R. (1991). *The laboratory of the mind: Thought experiments in the natural sciences*. Routledge.

Brown, J. R. (1994). *Smoke and mirrors: How science reflects reality*. Routledge.

Brown, J. R. (2004). Why thought experiments transcend empiricism. In C. Hitchcock (Ed.), *Contemporary debates in philosophy of science* (pp. 23–43). Blackwell.

Brown, J. R., & Fehige, Y. (2023). Thought experiments. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 ed.). Stanford University. https://plato.stanford.edu/archives/win2023/entries/thought-experiment/

Bryant, A. (2017). Keep the chickens cooped: The epistemic inadequacy of free-range metaphysics. *Synthese, 194*(7), 2581–2602. https://doi.org/10.1007/s11229-017-1398-8

Chakravartty, A. (2017). *Scientific ontology: Integrating naturalized metaphysics and voluntarist epistemology*. Oxford University Press.

Cooper, R. (2005). Thought experiments. *Metaphilosophy, 36*(3), 328–347.

Emery, N. (2025). Two Types of Naturalism and the Metaphysics of Science. *Análisis Filosófico*, *45*(Especial), 767-793.

Frappier, M., Meynell, L., & Brown, J. R. (Eds.). (2013). *Thought experiments in science, philosophy, and the arts*. Routledge.

Gähde, U. (2000). Gedankenexperimente in Erkenntnistheorie und Physik: Strukturelle Parallelen. In J. Nida-Rümelin (Ed.), *Rationalität, Realismus, Revision* (pp. 457–464). De Gruyter.

Gendler, T. S. (2000). *Thought experiment: On the powers and limits of imaginary cases*. Garland Press.

Goffi, S., & Roux, S. (2018). A dialectical account of thought experiments. In M. T. Stuart, Y. Fehige, & J. R. Brown (Eds.), *The Routledge companion to thought experiments* (pp. 439–453). Routledge.

Gramelsberger, G., & Feichter, J. (2011). Introduction to the volume. In G. Gramelsberger & J. Feichter (Eds.), *Climate change and policy: The calculability of climate change and the challenge of uncertainty* (pp. 1–8). Springer.

Humphreys, P. (2013). Scientific ontology and speculative ontology. In D. Ross, J. Ladyman, & H. Kincaid (Eds.), *Scientific metaphysics* (pp. 51–78). Oxford University Press.

Jaksland, R. (2023a). A trilemma for naturalized metaphysics. *Ratio*, 36(1), 1–10. https://doi.org/10.1111/rati.12344

Jaksland, R. (2023b). Naturalized metaphysics or displacing metaphysicians to save metaphysics. *Synthese*, 201, Article 199. https://doi.org/10.1007/s11229-023-04207-1

Ladyman, J. (2017). An apology for naturalized metaphysics. In M. H. Slater & Z. Yudell (Eds.), *Metaphysics and the philosophy of science: New essays* (pp. 141–162). Oxford University Press.

Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized*. Oxford University Press.

Lee, C. (2023). What analytic metaphysics can do for scientific metaphysics. *Ratio, 36*(3), 192–203. https://doi.org/10.1111/rati.12379

Melnyk, A. (2013). Can metaphysics be naturalized? And if so, how? In D. Ross, J. Ladyman, & H. Kincaid (Eds.), *Scientific metaphysics* (pp. 79–95). Oxford University Press.

Morgan, M. S. (2012). *The world in the model: How economists work and think*. Cambridge University Press.

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

Norton, J. D. (1991). Thought experiments in Einstein's work. In T. Horowitz & G. J. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 129–148). Rowman & Littlefield.

Norton, J. D. (1996). Are thought experiments just what you thought? *Canadian Journal of Philosophy, 26*(3), 333–366. https://doi.org/10.1080/00455091.1996.10717457

Okasha, S. (2018). *Agents and goals in evolution*. Oxford University Press.

Paul, L. A. (2012). Metaphysics as modeling: The handmaiden's tale. *Philosophical Studies, 160*(1), 1–29. https://doi.org/10.1007/s11098-012-9906-7

Ritchie, J. (2022). On the continuity of metaphysics with science: Some scepticism and some suggestions. *Metaphilosophy, 53*(2–3), 202–220. https://doi.org/10.1111/meta.12535

Schurz, G. (2020). Abduction as a method of inductive metaphysics. *Grazer Philosophische Studien*, 97(1), 1–25. https://doi.org/10.1163/18756735-000098

Sorensen, R. A. (1992). *Thought experiments*. Oxford University Press.

Stuart, M. T., Fehige, Y., & Brown, J. R. (Eds.). (2018). *The Routledge companion to thought experiments*. Routledge.

Suárez, M. (2024). The emergence of the modelling attitude. In *The Routledge Handbook of Philosophy of Scientific Modeling* (pp. 11-25). Routledge.

Tooley, M. (1977). The nature of laws. *Canadian journal of Philosophy*, *7*(4), 667-698.

Varga, S. (2021). The continuity of inquiry and normative philosophy of science. *Metaphilosophy*, *52*(5), 655-667.

Wilkes, K. V. (1988). *Real people: Personal identity without thought experiments*. Oxford University Press.

Williamson, T. (2016). Abductive philosophy. *The Philosophical Forum, 47*(3–4), 263–280. https://doi.org/10.1111/phil.12122

Winsberg, E. (2018). *Philosophy and climate science*. Cambridge University Press.

Weisberg, M., Okasha, S., & Mäki, U. (2011). Modeling in biology and economics. *Biology & Philosophy*, *26*(5), 613–615.