Beauty Leads to Truth: Aesthetic Induction on Consistency

Lyu FU[1]

Abstract

The belief that beauty leads to truth is prevalent among contemporary physicists. Far from being a private faith, it operates as a methodological guiding principle, essentially when physicists have to develop theories without new empirical data. However, it is unclear how beauty should be understood here for this belief to be justified not merely as useful but as true. In this article, I propose an interpretation of "beauty leads to truth" as "ugliness leads to falsehood," where "ugliness" refers to a lack of formal harmony, namely, a lack of consistency; in other words, "beauty leads to truth" is interpreted as "inconsistent theories cannot be true." As this article will show, this conviction (that inconsistent theories cannot be true) is indeed utilized as a methodological principle in scientific practice. Nevertheless, finding a justification is not easy, for this conviction is not merely a logical requirement, nor is it readily supported by direct observation or theoretical considerations. The sole non-circular justification seems to lie in a meta-induction: historically, inconsistent theories are less successful than their consistent successors. This constitutes an aesthetic induction, for (in)consistency can be understood as an aesthetic property, at least within a hermeneutic context, and it may perform a genuinely aesthetic role in this meta-induction. In this sense, "inconsistent theories cannot be true" is a specific instance of "beauty leads to truth," or, alternatively, "ugliness leads to falsehood."

Key words: mathematical beauty, paraconsistent logic, meta-induction, Dialetheism, Pythagoreanism, applicability of mathematics

Funding Statement

The author is supported by a doctoral scholarship from the China Scholarship Council (CSC). No additional funding was received for this specific study.

Data Availability

No new data were created or analyzed in this study. Data sharing is therefore not applicable.

[1] ✉Lyu.Fu@etu.univ-paris1.fr
https://orcid.org/0009-0003-6774-5998
IHPST (CNRS and Univ. of Paris 1 Panthéon-Sorbonne), UMR 8590, 13, rue Du Four, 75006 Paris, France

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Introduction

In this article, I focus on a belief that is not uncommon among contemporary physicists: that the mathematical beauty (of theories) can lead to physical truth. I encapsulate this belief as "beauty leads to truth." An important question is, then, how "mathematical beauty" should be understood here, or whether there exists an interpretation of "mathematical beauty" that would make this belief true. To address this, I begin in Section 1 with a brief overview of relevant historical background, followed in Section 2 by a summary of how this belief has been treated in the philosophical literature as a methodological principle. In Section 3, I propose interpreting "beauty leads to truth" as "ugliness leads to falsehood," where "ugliness" denotes a lack of formal harmony—namely, inconsistency. This yields the claim that "inconsistent theories cannot be true," a position widely held by physicists. Section 4 considers four possible justifications for this claim, and shows that only the historical justification, based on a meta-induction, avoids circularity. Sections 5 and 6 assess the extent to which this meta-induction qualifies as an aesthetic induction. I argue that (in)consistency is, at minimum, a hermeneutic aesthetic property and that it may play a genuinely aesthetic role in the meta-induction. However, this does not imply that the aesthetic preference for consistency was itself historically shaped by this meta-induction. Section 7 addresses several objections.

1. Historical Background

McAllister observes that "[o]ne of the most remarkable features of modern science is the conviction of many scientists that their aesthetic sense can lead them to the truth" (1999, p. 90). This statement refers solely to scientists' belief in the guiding role of aesthetics, without addressing whether this belief is justified. It can be further specified with a focus not on beauty but on mathematical beauty. To support this more focused claim, we can refer to testimonies from prominent scientists:

> Heisenberg (1971, p. 68): "If nature leads us to mathematical forms of great simplicity and beauty—by forms I am referring to coherent systems of hypotheses, axioms, etc.—[…] we cannot help thinking that they are 'true,' that they reveal a genuine feature of nature."

> Dirac (1970, p. 29): “A theory with mathematical beauty is more likely to be correct than an ugly one that fits some experimental data.”

> Penrose (1974, p. 267): “It is a mysterious thing in fact how something which looks attractive may have a better chance of being true than something which looks ugly.”

When and why did this belief arise? Bangu (2006) demonstrates that the beginning of the quantum era was a critical historical juncture:

> "At the beginning of the quantum era, physics faced a ‘grave crisis of ideas,’ since the traditional tools employed in theory construction turned out not to be transferable to the microworld phenomena. Experimentation and manipulation in (sub)atomic physics were often at an impasse too, for obvious reasons. An interesting general suggestion was made as to how to advance the research, Bohr’s complementarity principle. But, as Heisenberg pointed out, this guidance soon became unreliable, and Dirac came up with his Pythagorean, ‘indirect way’: develop the mathematical formalism in the direction indicated by your sense of beauty (‘perfect’ the formalism, as Dirac says), and then hope that it will guide your physical theory in the right direction. This happened indeed and, more often than not, the method worked well. Mathematical beauty served as a guide to physical truth indeed; moreover, the context was such that nothing else could help—not much help from the phenomenology of physics was available.” (p. 410–411)

A key point here is that, according to Bangu, this conviction that beauty leads to truth was not merely a private faith but a methodological principle at the beginning of the quantum era. Then, is it still one today? McDonnell (2017) argues that it is. She suggests that contemporary physicists employ two different research strategies:

> (S1): “Use models to extrapolate from known phenomena to unknown phenomena; testing the boundaries but keeping close to experimentally verifiable phenomena. Here the focus is on developing theoretical and experimental tools in tandem to discover new phenomena.”

> (S2): “Broaden and deepen the conceptual basis of mathematically beautiful theories to eliminate the known ugly bits and to encompass more phenomena. Here the focus is on developing a unified basis for all of physics; resolving the theoretical contradictions whilst retaining beautiful features.”

Regarding the second one, she emphasizes: “[in] applying (S2), experimental verification is secondary. It is even permissible to work on theories which contradict experimental data if they contain potentially interesting ideas” (Ibid., p. 21). String theory is considered a typical example of this strategy (*Ibid.*, p. 28).

Strategy (S2) may be justified by its background. As Hossenfelder (2018) notes, experimental verification in physics requires increasingly extensive time, compelling physicists to develop theories in the absence of new data. Consequently, mathematical beauty has (once again) become an essential guiding principle.

2. Philosophical discussion

Philosophical discussions focus more on the contexts and ways in which this belief is employed as a methodological principle.

Following discussions on the applicability of mathematics, I recognize a distinction between the context of discovery and the context of justification. To clarify their difference, a distinction between *information* and *inspiration* will be needed: the former allows us to rule out possibilities that have already been identified, while the latter enables us to conceive of possibilities not previously realized (see below Figure 1 Information and inspiration).

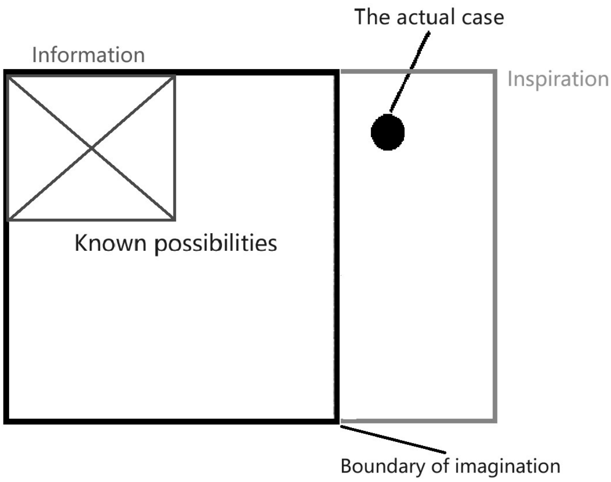

Figure 1 Information and inspiration

With this distinction, we can define the context of discovery as the context in which inspiration is sought to propose new hypotheses, and the context of justification as the context in which information is sought to test existing hypotheses (to judge whether they are true or false).

So, is "beauty leads to truth" applied as a methodological principle in the context of discovery or in the context of justification?

McAllister (1999) focuses on the context of justification. He argues that since experimental data is fallible, relying solely on it to test theories may lead to false negatives (deeming a correct theory false) and false positives (deeming an incorrect theory true). Beliefs such as "beauty leads to truth" help to mitigate these false negatives and positives (*Ibid.,* p. 92~95).

Moreover, as Hossenfelder (2018) demonstrates, such beliefs are invoked to justify theory choices under empirical underdetermination. When two theories are equally supported by empirical data yet make different predictions, scientists appeal to aesthetic considerations to defend their preference.

How is this principle applied in the context of justification? McAllister (1999) addresses this through the concept of aesthetic induction:

> "If a theory possessing an aesthetic property P scores notable empirical success, the community comes to regard P with increased favor and to expect future theories showing P to be successful too. On the other hand, if there later arise theories that lack P but are empirically more successful than the P-bearing theories, then the community's preference for future theories to show P wanes. " (p. 78)

More specifically, if past physical theories with mathematical beauty have achieved empirical success, then a new physical theory with mathematical beauty should also be expected to achieve empirical success.

This process is a form of meta-induction. Induction concerns phenomena: for example, the first raven is black, the second raven is black... thus, the next raven is likely to be black. Meta-induction concerns theories: for instance, in the case of pessimistic meta-induction (Laudan 1981), the first scientific theory turned out to be false, the second scientific theory turned out to be false... hence, the next scientific theory is also likely to be false. (For more on meta-induction, see Schurz 2008, where the definition is slightly different.)

Different from McAllister (1999), Bangu (2006), like Steiner (1989, 1995, 1998), explicitly claims to focus on the context of discovery (p. 394). In his description of the early days of the quantum era, he claims that "beauty leads to truth," as a methodological principle, was employed to address the "grave crisis of ideas" and played a role in “theory construction.” It is reasonable to assume that, in their discussions, this principle aids in the search for inspiration rather than information.

How, then, can this principle be applied in the context of discoveries? In other words, how does this principle help scientists discover hypotheses that had previously not been conceived of, or were even unconceivable? Bangu does not provide a detailed answer here, but we can turn to Steiner, whose answer lies in mathematical analogies[2]:

---

[2] To be fair, Steiner is not the first to express a similar idea; Poincaré (1905) had already pointed out that one of mathematics' primary contributions to physics is that mathematical analogy allows us to

> "My answer: by mathematical analogy. […] Mathematics itself thus provided the framework for guessing the laws of the atomic world, by providing its own classificatory schemes" (Steiner 1998, pp. 3–4).

I do not intend here to delve into the details of Steiner's account of mathematical analogies (see Steiner 1989 and 1998 for further details); rather, I am content to roughly illustrate how, from his perspective, mathematical analogies aid scientists in 'guessing' new hypotheses. In brief, new physical hypotheses arise from a guess-before-understanding approach: first, one hypothesizes that the mathematical structure underlying a phenomenon resembles certain mathematical structures in existing theories (i.e., one makes a mathematical analogy), and then attempts to assign physical significance to the guessed mathematical structure. According to Steiner, Schrödinger formulated his famous equation by hypothesizing certain mathematical similarities between the mechanics of microscopic particles and the equations of wave optics; the physical picture implied by this wave mechanics—markedly different from classical mechanics—emerged then from the interpretation of this equation (1998, pp. 79–82; for critiques, see Bueno & French 2018, Chapter 1.)

In this process, what role does beauty play? Unlike Bangu, who seems to think that not all mathematics is beautiful, Steiner emphasizes the intrinsic connection between mathematical structure and beauty. He argues that "concepts are selected as mathematical because they foster beautiful theorems and beautiful theories" (1998, p. 64). Therefore, for him, mathematical analogy is itself a practice grounded in aesthetic considerations.

It is worth noting that, in Steiner and Bangu's accounts, mathematics—or beautiful mathematics—not only helps scientists expand the boundaries of their imagination to conceive new hypotheses but also, more surprisingly, directs this expansion in a correct direction. This belief implies that hypotheses inspired by beautiful mathematics are more likely to be true than those inspired by ugly mathematics—thus, implying a (probabilistic) judgment regarding the truth of a hypothesis. In this sense, their discussion touches not only on the context of discovery but also on the context of justification. Therefore, in the following discussion, I will focus on the role of the belief in the context of justification.

3. Justifiable new interpretation

How can the use of "beauty leads to truth" as a methodological principle in the context of justification be itself justified?

---

generalize empirical observations effectively (pp. 157–158). However, Steiner more explicitly emphasizes the role of mathematics in the context of discovery.

McAllister (1999) contends that it can only be *pragmatically justified*:

> "There may or may not be correlations between theories' having particular aesthetic properties and their having high degrees of empirical adequacy. If there are no such correlations, then no method of forming criteria for theory evaluation will identify any. But if some such correlations exist, then inductive projection will be at least as likely to discover them as any alternative procedure for formulating criteria. " (p. 101)

Simply put, this method is a "no-lose" strategy—it does no harm. He does not believe that we have genuinely identified any aesthetic properties that have a real connection to truth:

> "Contrary to Dirac, Einstein, and others, I see little evidence that aesthetic properties correlated with high degrees of empirical adequacy in theories have yet been identified in any branch of science. " (p. 102)

However, this pragmatic justification is not valid, especially in cases of theory choices (under empirical underdetermination), for aesthetic considerations may lead theoretical research “astray” (Hossenfelder, 2018). In other words, the methodological principle “beauty leads to truth” is *not* a no-lose strategy. Therefore, in order to justify the application of this belief in theory choices, we need an alternative line of justification.

I argue—contra McAllister—that there may indeed be an aesthetic property genuinely related to empirical adequacy: consistency. Its connection to truth is better captured by the formulation "ugliness leads to falsehood" rather than "beauty leads to truth." In other words, if a physical theory is mathematically inconsistent, then it cannot be true. This principle is, in fact, employed in theory choices, and I will show that—despite the apparent triviality—it cannot be easily justified, and may indeed involve consistency as an aesthetic property.

Let us start with an example shared by McAllister (1999), Bangu (2006), and McDonnell (2017): Dirac’s rejection of quantum electrodynamics (QED) based on aesthetic judgment:

> "When in 1950 Dyson asked Dirac what he thought of the new development in quantum electrodynamics, he answered, ‘I might have thought that the new ideas were correct if they had not been so ugly.’" (Kragh 1990, p. 184)

Clearly, this example concerns not theory construction but theory testing. Dirac is judging the potential correctness of a theory based on his aesthetic sense, which involves the context of justification rather than that of discovery. But why did he consider the QED ugly? Dirac stated:

> "Most physicists are very satisfied with the situation. They say: 'Quantum electrodynamics is a good theory, and we do not have to worry about it any more.' I must say that I am very dissatisfied with the situation, because this so-called 'good theory' does involve neglecting infinities which appear in its equations, neglecting them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves neglecting a quantity when it is small—not neglecting it just because it is infinitely great and you do not want it! " (quoted in Kragh 1990, p. 184)

Here, I will not delve into the technical details; it suffices to say that this issue is related to renormalization. Like Dirac, Feynman (1985) also expressed doubts about some of the mathematics in QED:

> "The shell game that we play […] is technically called 'renormalization.' But no matter how clever the word, it is what I would call a dippy process! Having to resort to such hocus-pocus has prevented us from proving that the theory of quantum electrodynamics is mathematically self-consistent. It's surprising that the theory still hasn't been proved self-consistent one way or the other by now; I suspect that renormalization is not mathematically legitimate. " (p. 128)

This observation of Feynman gives us a reason to say the ugliness Dirac felt was due to a concern of inconsistency. A theory is inconsistent if and only if it is contradictory, that is to say, it entails a statement and its negation.

Then this case can be viewed as an illustration of a specific instance of "beauty leads to truth": "ugliness leads to falsehood," where "ugliness" refers to formal inconsistency. Here, we observe a belief shared by many, if not most, physicists and philosophers of science: if a physical theory[3] is mathematically inconsistent (in other words, if it contains contradictions), then it cannot be true. Furthermore, as a corollary, if two physical theories are mathematically inconsistent with each other, they cannot both be true (since the combined theory, being inconsistent, cannot be true).

This belief can be traced back at least to Leibniz, whose "principle of necessary things" asserts that "whatever implies a contradiction is false" (Leibniz 1989, p. 19). More recently, W. Frank (2005), while talking about physical theories, argues that "contradictions imply serious error" and that "nature cannot realize contradictions" (p. 857). More importantly, much scientific practice is conducted under the guidance of this belief. Weinberg's rejection of his nonlinear quantum theory is an example:

[3] A theory is usually considered to be deductively closed, meaning that all its deductions are contained within the theory itself. Of course, deduction can be defined within different logics. However, some scholars argue that deductive closure should be abandoned (Smith 1998a, 1998b). Here, I do not intend to enter this debate. It suffices to say that broader definitions of a theory, such as one that considers a theory to be simply a set of statements, can be adopted.

considering that the linear nature of quantum mechanics prevents quantum systems from exhibiting chaotic behavior, he proposed, "a slightly nonlinear alternative to quantum mechanics." Even though subsequent experiments failed to detect the nonlinear effects predicted by this alternative, Weinberg was not overly disappointed, as “there was no reason to believe that these corrections should be just large enough to show up in the first round of experiments designed to search for them.” What truly disappointed him was that the theory faced certain “purely theoretical internal difficulties,” specifically its conflict with special relativity: “the nonlinearities of the generalized theory could be used to send signals instantaneously over large distances, a result forbidden by special relativity”(Weinberg, 2011, pp. 87–89). Arguably, Weinberg’s reason for abandoning his nonlinear theory was the belief that two contradictory theories cannot both be true, and special relativity was unlikely to be incorrect. If the belief that inconsistent theories cannot be true cannot be itself justified, then scientific practices based on this belief would also lack justification—a clearly serious consequence.

4. Four Justifications

So how might the conviction that “inconsistent theories cannot be true” be justified? In this section, I will examine four possible defenses, only the last of which—historical justification—proves plausible. In the following sections, I will show to what extent the meta-induction at the heart of this justification qualifies as an aesthetic induction. Taken together, they lay the groundwork for this paper’s core claim : that there is a meta-inductively justified link between truth and (in)consistency as an aesthetic property.

The first justification comes from logic. In logic, there is a principle known as the principle of explosion: *ex contradictione sequitur quodlibet* (from a contradiction, one can derive any proposition). Thus, if we derive from a contradictory theory (containing p and $\neg p$) that, say, ravens are black, we could also derive that ravens are white (denoted as q) from, for example, the two following valid inferences:

$$\frac{p}{p \vee q}$$

$$\frac{\neg p \quad p \vee q}{q}$$

However, this would render the theory both false and useless. Thus, consistency may be regarded as a reasonable logical constraint on scientific theories.

Physicists and philosophers of science have indeed used similar arguments. For example, Heisenberg argued that “we can scarcely describe nature without having something consistent” (Heisenberg in AHQP, February 25, 1963). Why? He explained:

> "An inconsistency can never be made to work. I remember that at that time we frequently quoted this sentence from Hilbert, who had studied the inconsistencies of mathematical axioms and who had proved that if the mathematical system of axioms does contain an inconsistency -A equal to non A- then you can prove everything from this scheme. That we applied, of course, with great pleasure to physics. As soon as you came into a real inconsistency, then you could go anywhere. That, of course, is nowhere. Anywhere, nowhere, it's just an accident. " (Heisenberg in AHQP February 25, 1963)

Similarly, Popper argued that "[the acceptance of inconsistency] would mean the complete breakdown of science. This can be shown by proving that if two contradictory sentences are admitted, any sentence whatsoever must be admitted" (1940, p. 408).

However, this justification does not hold up under scrutiny. First, the principle of explosion is not always invoked[4], and, more fundamentally, it can be discarded, as in *paraconsistent logic*. According to the Jaśkowski-da Costa definitions:

- A theory is inconsistent if it can derive both a statement and its negation.
- A theory is trivial if it can derive all statements.
- A theory is *paraconsistent* if it is inconsistent but not trivial.
- *Paraconsistent logic* is logic designed to handle paraconsistent theories (for more on paraconsistent logic, see Arruda 1980 and da Costa & Marconi 1987).

In fact, many scientific theories, such as Newtonian cosmology (Norton 1993), the old quantum theory, including blackbody radiation theory (see Norton 1987), and Bohr's theory (see Brown 1993), do contain contradictions, yet some of them (such as Bohr's theory) still achieve empirical success. One application of paraconsistent logic is to help understand how such contradictory theories are and can be controlled (see Da Costa & French, 2003, Chapter 5). Thus, consistency is not a logical requirement that scientific theories must meet, for they can be controlled in a paraconsistent way.

Second, this justification appears question-begging. The principle of explosion is a consequence of the absence of models that can satisfy a contradictory set of propositions. Generally, a semantics for a logic $\mathcal{L}$ can be defined as a set W of objects with a function m from the set of formulas into the power set of W. Given an element w of W and a formula f, w is a model of f (meaning w satisfies f, or that f is true in w) iff $w \in m(f)$. A model of a set of formulas is a model of each of its components. A formula φ is a semantic consequence of a set of formulas P if and only if every model of P is a model of φ, that is, $m(P) \subseteq m(\varphi)$.

---

[4] If we adopt a broader definition, viewing a "theory" as a set of statements that does not need to be deductively closed, then the logical justification fails directly.

Now, consider a contradictory set of formulas, such as $\{p, \neg p\}$. Suppose that no model can satisfy this set; in other words, no model can make p and $\neg p$ simultaneously true, so $m(\{p, \neg p\}) = \emptyset$. Since $\emptyset$ is a subset of any set, for any proposition q, we have $m(\{p, \neg p\}) \subseteq m(q)$, making q a consequence of $\{p, \neg p\}$. If $\mathcal{L}$ is complete and sound for W, it will allow us to derive any proposition from a contradiction, thus giving us the principle of explosion. Conversely, if some elements in W are models of $\{p, \neg p\}$—in other words, if m({p, ¬p}) is not empty—then only propositions that are true in all models of $\{p, \neg p\}$ can be derived from $\{p, \neg p\}$. If not all propositions are always true in these models (in which $\{p, \neg p\}$ is true), the principle of explosion does not hold. Therefore, the principle of explosion is a consequence of the absence of models that satisfy contradictory formulas. Using it to argue that no model can satisfy a contradictory set of formulas (which, in our context, means no physical phenomenon can make a contradictory set of propositions true) seems to constitute a circular reasoning.

This leads us to the second justification, which I call empirical justification: in fact, we have not found any directly observable physical phenomenon that could make a contradictory theory true, and this gives us good reason to expect that, in W. Frank's words, "nature cannot realize contradictions."

However, this justification also does not hold up under scrutiny. First, an inconsistent theory does not necessarily produce contradictory *observable* predictions. A typical example is what Brown calls old quantum theory, which is primarily the Bohr model; in this model, energy is inconsistently assumed to be both discrete and continuous (Brown 1993). Despite formal contradictions, old quantum theory effectively avoids contradictory observable physical predictions. According to Brown, this is due to a contextualization strategy:

> "As it developed, old quantum theory came to include explicit rules for the application of the conflicting principles involved. Thus, while in any context of application incompatible principles would be invoked, there were identifiable subcontexts in which they were isolated, preventing the conjunctive application of contrary claims at any point. It is within these subcontexts that we can treat the claims involved in such theories as true" (Brown, 1990, p. 284; for critiques of this approach, see the local-consistency objection in section 7).

Thus, the absence of direct observation of contradictory phenomena (that is, phenomena that can make a contradictory set of statements true) does not demonstrate that inconsistent theories cannot be true.

If the issues with empirical justification were limited to this, it would allow us to support a weaker assertion: a theory which contains directly observable contradictions cannot be true. This weaker assertion, with its corollary that two theories cannot both be true if they are mutually contradictory at the empirical level, would still justify

Weinberg's abandonment of his nonlinear theory, along with similar practices. However, even this weaker assertion is not well justified, as the physical significance of a formal contradiction depends on interpretation. Contradiction seems to be an attribute of theories rather than of phenomena. Therefore, instead of saying we have not observed any contradictions in nature, it would be more accurate to say we have chosen not to represent observed natural phenomena with contradictory sets of propositions. However, how can this choice be justified *so that we can claim inconsistent theories cannot be true*? Indeed, the choice itself can be defended—for example, by appealing to the simplicity of consistent theories—but if that's the case, inconsistent theories would merely be more complex, not untrue. It's unclear how this justification could stand without being question-begging. Moreover, there are indeed scholars who attempt to use contradictory sets of propositions to provide a more direct representation of natural phenomena. A notable example is found in da Costa&De Ronde (2013), where the authors try to account for the contradictory properties of quantum superpositions within an explicitly inconsistent theoretical framework.

The third justification is physical justification, which argues that we have good reasons to accept the speculation that a consistent final theory or theory of everything (TOE) exists. Then if a theory (or set of theories) is inconsistent, it must contain elements incompatible with the TOE, and thus this set of theories cannot be strictly true. The idea behind this justification is that "inconsistent theories cannot be true" is a corollary of the speculation of the existence of a consistent TOE, which itself appears to be defensible or, at least, a good thing to accept (for discussions on whether the existence of a TOE is justified or desirable to accept, see Achinstein 2018, Chapter 5).

This physical justification also faces significant issues. The key question is: why should the final theory or TOE be consistent? On this point, I found only one justification in literature that *seems* independent of our question: a final theory is not only a theory that *cannot* be explained by deeper principles, but also one that *does not need* any further explanation. It *must be* as it is. Any slight modification would destroy it (Achinstein 2018, pp. 227–229). However, the problem is that it is almost impossible for this theory to be, in Weinberg's words, "logically inevitable". So, what properties should it have?

> "In my view, our best hope along this line is to show that the final theory, though not logically inevitable, is logically isolated. That is, it may turn out that, although we shall always be able to imagine other theories that are totally different from the true final theory (like the boring world of particles governed by Newtonian mechanics), the final theory we discover is so rigid that there is no way to modify it by a small amount without the theory leading to logical absurdities. In a logically isolated theory, every constant of nature could be calculated from first principles; a small change in the value of any constant would destroy the consistency of the theory." (Weinberg

2011, p. 236).

In other words, along this line of argumentation, the final theory would not only be consistent but also isolatedly consistent. However, why should isolated consistency mean that the final theory no longer requires explanation and must be as it is? The reasoning behind this seems to rely on the idea that an inconsistent theory cannot be true, so that any minor change would render the theory false, making it what it is. Thus, this argument still appears to risk being question-begging.

The fourth and final justification is historical justification: in the history of science, inconsistent theories have not been as successful as their consistent counterparts (just as the Bohr model was less successful than quantum mechanics in accurately predicting atomic behavior).[5] Therefore, by meta-induction, consistent theories can be expected to be more successful than their inconsistent alternatives. A complete, true theory of the universe would undoubtedly be the most successful theory and, therefore, must be consistent.

There are similar observations in the literature. Post (1971) considers mathematical inconsistencies as a kind of "formal flaw" with heuristic power, suggesting that "a formal 'flaw' in the old theory, a neuralgic point, is the starting point for the new." Similarly, Da Costa *et al.* (2003, p. 91) points out that "inconsistent theories in science typically point the way to a consistent successor." Building on this observation, the historical justification adds that what Post calls "the new [theory]" and what Da Costa *et al.* call the "consistent successor" have achieved greater success. This fact, by meta-induction, justifies the belief that "inconsistent theories cannot be true."

The historical justification, of course, is open to criticism from the history of science, but it is at least not as overtly circular as the other justifications, and its core claim (that inconsistent theories are less successful than their consistent successors) is at least apparently plausible. Thus, I delegate the task of criticizing this justification to those who wish to challenge it.

5. Explanatory Power Argument

If historical justification is valid, then there exists a meta-inductively justified link between (in)consistency and empirical success: the first consistent theory was more successful than its inconsistent counterpart, the second likewise, ... therefore, the next consistent theory will be more successful too. In this section, I argue this constitutes an aesthetic induction. But what exactly does that mean?

Indeed, the concept of "aesthetic induction" is ambiguous. Firstly, it admits at least

[5] As will be discussed in Section 6, it may be that "inconsistent theories," as referenced in this justification, should be understood to mean "theories that *turn out* to be inconsistent." If so, theories whose constitutive representations are deliberately inconsistent could still be true.

two interpretations, each corresponding to a distinct question: (1) How are aesthetic preferences formed? (2) How can the belief that beauty leads truth be justified? As an answer to the former, aesthetic induction is understood as a *historical process*, sometimes regarded as purely psychological, involving the mere-exposure effect (Kuipers 2002, Ivanova 2017), and sometimes as cognitively rich, involving epistemic engagement (Currie 2023). As an answer to the latter, aesthetic induction functions as a *retrospective justification*, essentially an inference. It depends on the relevant property—here, (in)consistency—being aesthetically related but provides no further explanation; it does not even require scientists to find consistency more beautiful than an inconsistency. Moreover, "aesthetically related" can itself be further divided into genuine relevance and superficial relevance, and the induction may or may not rely on this aesthetic relation. Reflecting this ambiguity, I will discuss four distinct theses in what follows:

> **Hermeneutic Thesis**: (In)consistency is an aesthetic property *in a hermeneutic sense*[6]. In other words, within a hermeneutic context, it is plausible to claim that consistent theories, precisely due to their consistency, are more beautiful than their inconsistent counterparts.
>
> **Intrinsicness Thesis**: This meta-induction intrinsically involves (in)consistency *as an aesthetic property*. That is, the link between (in)consistency and empirical success is mediated and regulated by a specific aesthetic value.
>
> **Genuineness Thesis**: (In)consistency elicits *genuine* aesthetic responses. In other words, the aesthetic properties and values discussed above are not merely epistemic properties or values in disguise.
>
> **Formation Thesis**: The aesthetic preference for consistency is historically formed by the meta-induction as a historical process.

I primarily support the Hermeneutic Thesis. I think it should be accepted unless we get a better but incompatible hermeneutic account. The Intrinsicness Thesis and Genuineness Thesis require additional empirical support, but I think we have reasons to stay optimistic. As for the Formation Thesis, I consider it no more persuasive than its rivals. In this section, I defend the Hermeneutic Thesis. The remaining three will be discussed in the next section.

Hermeneutic Thesis. This thesis addresses a critical hermeneutic task: when a

[6] In a hermeneutic context, we seek to discern what someone means when they call something beautiful or ugly. Crucially, such utterances are not reason-based but cause-based: an object is experienced as beautiful not because it fits a preconceived concept or definition of beauty, but because it elicits aesthetic pleasure (Hilgers, 2016, pp. 18–19). The hermeneutic inquiry thus becomes: which feature of the object provokes this response? That feature—identified as responsible for the aesthetic reaction—is an aesthetic property in the hermeneutic sense.

scientist judges a theory as (potentially) true or false based on aesthetic considerations, which property of the theory is responsible for her aesthetic utterance—for example, "This theory is ugly"? Moreover, how can her judgment ("This theory is too ugly to be true") be justified and criticized? The Hermeneutic Thesis provides *an* answer to these questions: (in)consistency may be responsible for her aesthetic statement; the connection between (in)consistency and empirical success is justified by a meta-induction; we can criticize her judgment by pointing out, for instance, that the theory is not actually inconsistent.

Two points deserve attention. First, following McAllister (1996), I adopt a projectivist view, which distinguishes aesthetic *values* from aesthetic *properties*. In this framework, (in)consistency is treated not as an aesthetic value but as an aesthetic property. Aesthetic values arise from aesthetic evaluations, whereas aesthetic properties are features that give rise to such evaluations; their verification is independent of aesthetic evaluations. Inconsistency is not a kind of ugliness; rather, it is what makes a theory ugly in a certain way. Second, the Hermeneutic Thesis does not require a *genuine* aesthetic evaluation behind the scientist's aesthetic utterance. It is thus compatible with Todd's (2008) provocative claim that aesthetic judgments are merely "masked" epistemic assessments.

The plausibility of the Hermeneutic Thesis hinges on the extent to which (in)consistency can explain or account for aesthetic utterances. My strategy for defending it is to support a broader claim that a theory's degree of consistency (partly) explains or accounts for its degree of beauty. To this extent, consistency and inconsistency can be viewed as responsible for aesthetic utterances and are thus aesthetic properties in the hermeneutic sense. This is what I call the explanatory power argument. It is based on an account of mathematical beauty in scientific theories, whose core claims are: (1) consistency admits degrees, with inconsistency being a segment of the spectrum; and (2) the aesthetic value of a theory depends partly on its degree of consistency. Call this the consistency account (see below Figure 2 Consistency Account of Theoretical Beauty).

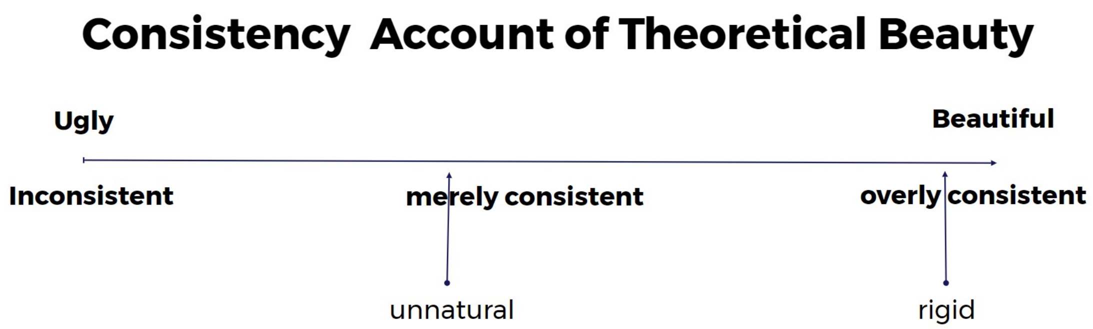

Figure 2 Consistency Account of Theoretical Beauty

The key to measuring the degree of consistency lies in the distinction between constitutive and derivative representations, where the former provides confirmability condition for the latter (Pincock 2012, p. 135). A classic example is Newton's laws of motion and the law of universal gravitation:

> "Within the context of Newtonian physics, [...] the only way in which we know how to give empirical meaning and application to the law of universal gravitation is by presupposing that the laws of motion are true: if the latter principles are not true (in the sense that there exists a frame of reference in which they hold) then the question of the empirical truth (or falsity) of the law of universal gravitation cannot even arise." (Friedman 2001, p. 74)

In simpler terms, the confirmable content, which is *a posteriori*, of the law of universal gravitation has to be identified through the laws of motion, which are regarded as "relativized a *priori*" (Pincock, 2012, p. 136). According to the logical relation between *a posteriori* statements and their *a priori* constitutive representation, consistent theories range from *merely consistent* to *overly consistent*: if the *a posteriori* statements are not only compatible with the constitutive representation but are also its logical consequences—in other words, if slight modifications to these statements lead to contradictions—then the theory is overly consistent. Conversely, if these statements are independent from the constitutive representation, the theory is merely consistent.[7] Finally, if these statements contradict the constitutive representation, the theory is inconsistent. Note that in the consistency account, "inconsistency" denotes a mismatch between a theory's constitutive framework and its empirical content. Thus, according to this account, the inconsistency provokes aesthetic displeasure *only if it lies between a theory's form and its content*. This is important, and I shall return to it in the next section.

The consistency account helps clarify aesthetic properties commonly discussed in physics. Two examples suffice here. The first is "rigidity". For instance, modifying the exponent in Newton's law of gravitation from -2 to -2.01 would not cause contradictions in Newtonian mechanics. However, in relativity theory, such a change violates "underlying assumptions of the theory" (Weinberg 2011, p. 105). This shows, according to Weinberg, that relativity theory is more "rigid" than Newtonian mechanics (*Ibid.,* p. 105–106). Rigidity is regarded as an aesthetic property (*Ibid.,* p. 149–150);this view is shared by some other physicists (see Hossenfelder 2018). As Weinberg acknowledges, the concept of rigidity is "difficult to pin down at all precisely" (*Ibid.,* p. 106). Nevertheless, the consistency account provides at least an initial clarification: the closer a theory is to over-consistency, the more rigid it becomes.

The second example, slightly more complicated, is (un)naturalness. For instance,

---

[7] To avoid misunderstanding, let me clearly state that I do not claim overly consistent theories are more likely to be true than merely consistent ones.

within the Standard Model, the mass of the Higgs particle is generally regarded as "unnatural," because obtaining the correct mass requires fine-tuning the bare mass and quantum corrections to be extremely close yet not identical. This appears highly artificial due to the extremely low probability that two parameters could acquire such closely matching values (see Barbieri & Giudice 1988 for a quantitative criterion). According to the consistency account, the objectively identifiable feature responsible for (un)naturalness is the mere-consistency of the theoretical framework with respect to certain measured parameters—in other words, the fact that these parameters can vary within a certain range without undermining the internal consistency of the theory. Since the theoretical framework exhibits no preference for values within the range, the parameters are given a uniform distribution as *a prior* probability. As these parameters are supposed, in this theoretical framework, to be independent, the probability that two of them take such closely aligned values is extremely low, which makes the theory appear artificial and thus "unnatural." In short, unnaturalness can be reduced to a mere-consistency or a low degree of consistency, where the degree can be precisely quantified in terms of probability.

Considering the explanatory power of (in)consistency for the aesthetic evaluation of physical theories, I claim that we have good reason to accept it as aesthetic properties—unless we discover a better yet incompatible hermeneutic model.[8]

6. Intrinsicness, Genuineness and Formation

Intrinsicness Thesis. The argument in Section 5 faces a challenge from what I call "job objection". According to this objection, consistency does a purely logical job in the meta-induction: even if (in)consistency evokes no aesthetic response, inconsistent theories still cannot be true. If this objection is valid, then what we have is a logical meta-induction rather than an aesthetic one. Then the Intrinsicness Thesis fails.

My response to this objection relies on two conjectures based upon the consistency account. According to this account, inconsistency is weaker than mere-consistency in precisely the same sense as mere-consistency is weaker than over-consistency. This is not trivial. It implies that inconsistency only evokes discomfort—leading to judgments of "ugliness"—when it is situated between constitutive and derivative representations. In other words, according to this account, if a theory is constructed in a deliberately paraconsistent, thus inconsistent, framework (constitutive representation), or more simply, if a theory is constitutively inconsistent, it will not provoke aesthetic displeasure: this is my first conjecture. As noted earlier, da Costa & De Ronde (2013) proposed an explicitly inconsistent theoretical framework for quantum superpositions. The entailed theory does not seem to evoke the typical aversion scientists display towards inconsistent theories. More importantly, the conviction that "inconsistent theories cannot be true" may not apply straightforwardly

[8] To be fair, I am not the first to claim that (in)consistency is an aesthetic property. Ivanova (2020) has already acknowledged it as an aesthetic property.

to this particular case. A similar example is Dirac's theory of quantum mechanics. In his framework, Dirac used mathematical "fictions" such as the statement that "every self-adjoint operator can be put in diagonal form", as well as "improper" functions like the delta function, which are self-contradictory (Von Neumann, 2018, p. 2; for a philosophical analysis, see Bueno 2005 or Bueno & French 2018, chap. 7). However, his theory does not seem to have been regarded as untrue because of these inconsistencies. Hence, my second conjecture is that theories which are *constitutively inconsistent* will not be directly judged as untrue, unlike other theories that *turn out to be inconsistent.*

If these two conjectures hold[9], then the Intrinsicness Thesis gains support. This thesis claims that the capacity of (in)consistency to evoke aesthetic responses is intrinsically linked to the meta-induction. In other words, the link between (in)consistency and empirical success is mediated by an aesthetic value.[10] The Intrinsicness Thesis allows us to introduce such a value and claim that (in)consistency is considered relevant to empirical success only to the extent that it influences this value. Which one? Historically, consistency has been associated with harmony[11], and harmony has been connected to the classical conception of beauty[12] (Sartwell 2024). Following this, I suggest this aesthetic value essentially involves a kind of harmony, or in Todd's (2008) term, "appropriateness", between form and content (see Murphy 2023, Section 3, for more on form and content in aesthetics). When scientists discover that a theory—or a combination of theories—is inconsistent, their first instinct is not to suspect that nature itself is contradictory, but rather that our theoretical framework is not well suited to nature. It is this mismatch between form and content that gives rise to the sense of ugliness. If consistency were doing a purely logical job, we would not be justified in introducing an aesthetic value as a regulating parameter; consequently, we should not expect scientists to display different epistemic attitudes toward theories that are constitutively inconsistent (or more precisely, theories that are constructed within deliberately inconsistent constitutive representation) versus those that turn out to be inconsistent.

Genuineness Thesis. (In)consistency has been viewed so far as an aesthetic property

[9] These two conjectures concern only scientists' reactions, without saying whether these reactions can be justified. Nevertheless, it is quite possible that the historical justification discussed in Section 4 should indeed be restricted to cases where theories turn out to be inconsistent.

[10] Note that the Intrinsicness Thesis itself does not assert that this value is genuinely aesthetic—it might still be a "masked" epistemic value.

[11] When Leibniz (1989, p.22) spoke of harmony, he seemed to refer to the consistency of propositions describing the world (including the assumption of God's existence). Similarly, Dedekind used the term "inner harmony (*innere Harmonie*)" to refer to consistency (Dedekind 1930–32, Vol. 3, p. 343, Preface to the third edition, published in 1911)

[12] Interestingly, mathematical beauty seems to be more broadly connected with the classical arts: among mathematicians who listen exclusively to classical music, 66% consider the beauty of mathematics more important than its utility, while only 33% of those who listen to other types of music share this view. Similarly, 51% of mathematicians who read only classical literature rank mathematical beauty above utility, compared to 29% among those who do not read classical literature (Menger & Verschueren, 2023, p. 61).

responsible for a certain aesthetic value. It is presupposed that the process from (in)consistency to an aesthetic utterance (e.g., "Beautiful!") involves *genuine* aesthetic evaluation. However, this presupposition can be questioned. As Todd (2008) points out, aesthetic utterances can be driven by non-aesthetic satisfaction. Someone might exclaim, "Beautiful!" simply because her favored sports team scored a goal. Todd further raises two reasons for doubt. First, there exists a "suspicious overlap" between epistemic properties and aesthetic properties. Thus, neither the output (the aesthetic utterance "Beautiful!") nor the input (the degree of consistency) definitively reveals whether the underlying process is aesthetic or epistemic. Moreover, it is hard to believe that the intermediate process is genuinely aesthetic, for genuine aesthetic evaluations are required to be disinterested, or, in other words, they should not involve (epistemic or practical) interests. The utterance "Beautiful!" should express disinterested pleasure in order to be genuinely aesthetic. However, given the functional nature of scientific activities, Todd argues that it's difficult to imagine these evaluations as entirely free from epistemic interests (Todd 2008, p. 67).

My response to this challenge is threefold.[13] First, take over-consistency (or rigidity) as an example. It lies beyond Todd's so-called "suspicious overlap." If the consistency account is correct, rigidity can indeed be regarded as an epistemic property. But it remains unclear whether it qualifies as an epistemic *virtue* (which evokes epistemic *satisfaction*). Given equal empirical support, it is doubtful that a more rigid theory would be judged epistemically superior. As Hossenfelder (2018, pp. 74–75) notes, rigidity may indicate a "dead end" in theory development. Thus, the burden of proof rests on those who argue that rigidity generates epistemic rather than aesthetic satisfaction. Second, evaluation can be disinterested in two senses. An evaluation is subjectively disinterested if it involves no *intentional consideration* of interests; it is objectively disinterested if the *formation* of the underlying preference (through natural selection or a psychological process) is not controlled by interests. Genuine aesthetic evaluation only requires subjective disinterestedness, not objective one.[14] Therefore, nothing prevents us from having genuine aesthetic evaluations in scientific and mathematical contexts. As Ivanova (2017) notes, this is supported by neuroscience research (Zeki *et al.*, 2014). Lastly, as Todd himself acknowledges, epistemic and aesthetic evaluations may not be strictly separable. Currie (2023a), for example, argues that aesthetic appreciation exhibits "partial sensitivity" to doxastic states. Certain stages of epistemic evaluation may evoke aesthetic responses, and epistemic and aesthetic pleasures may even coincide (see Turner 2019 for a systematically developed case study on the intertwining of aesthetics and epistemology). If this is true, the overlap between epistemic and aesthetic properties would not be surprising.

[13] Appealing to the harmony between form and content also helps address Todd's challenge, as it allows one to evaluate a theory's aesthetic value without relying on aesthetic utterances. Indeed, Murphy (2023) adopts this approach. I thank the anonymous reviewer for highlighting this point.

[14] Indeed, most discussions in aesthetic literature do not address objective disinterestedness. See Hilgers (2016) for a systematic review of aesthetic disinterestedness.

Given recent developments in neuroaesthetics, whether scientists' judgments involve epistemic or aesthetic satisfaction has become largely an empirical issue. Yet, since contemporary physicists themselves treat attributes like (un)naturalness and rigidity as aesthetic properties, I think we have reason to remain optimistic about the genuineness of physicists' aesthetic utterances until negative empirical evidence emerges.

Formation Thesis. Why do scientists exhibit aesthetic preferences toward consistent theories rather than inconsistent ones? How do they form such aesthetic preferences ? The first possible answer involves the mere-exposure effect: undoubtedly, theories that achieve greater empirical success receive more exposure (they appear in textbooks, popular science, etc.), and consistent theories have historically been more empirically successful than inconsistent ones. According to the mere-exposure effect, "mere repeated exposure of the individual to a stimulus object enhances his attitude toward it," where "enhance" means making the attitude more positive (Zajonc 1968). Consequently, consistent theories come to be perceived as more beautiful. The mere-exposure effect is regarded as an implicit response suggested by McAllister through his notion of aesthetic induction (Kuipers 2002; Ivanova 2017). Indeed, it has been empirically verified that mere exposure shapes aesthetic preferences (Cutting 2003). An advantage of this account may be that it highlights an interaction between empirical success and aesthetic preference. However, as Ivanova (2017, 2020) points out, this answer remains incomplete. More recent empirical studies show that "mere exposure to bad paintings [...] decreases liking for them," indicating that "something other than mere exposure plays a role in judgments of paintings" (Meskin et al. 2013, p. 159).

There are two possible ways to address this difficulty. The first involves replacing the notion of exposure with a cognitively richer concept. I consider this as an attempt to rescue the Formation Thesis. Currie (2023a), with the emphasis on "epistemic engagement," constitutes an initial proposal.[15] According to Currie, an agent epistemically engaged in a knowledge-directed process will attune her aesthetic faculties for epistemic purposes. Nevertheless, it is unclear why this attunement should involve aesthetic *preferences*. Undoubtedly, epistemic engagement alters the agent's attentional mechanisms, making her more sensitive to epistemically significant details. However, given the limitations of the mere-exposure effect, this does not explain why she would regard such details as beautiful or ugly.[16]

[15] Currie's (2023a) approach challenges the claim that factivity or empirical success can *directly* shape aesthetic preferences. However, in our case, empirical success or factivity *indirectly* influences aesthetic preferences by affecting theories' exposure. Thus, the first response above remains immune to her criticism in this respect.

[16] To be fair, Currie (2023a) does give a solution, which resembles that in McAllister's (1999): paradigms, in Kuhn's sense, serve as exemplars of beauty—but why should this be the case? And isn't this precisely the phenomenon we are attempting to explain here? I believe a different answer is required here. See Currie (2023b) for more on his view of the formation of aesthetic preferences.

The second strategy abandons partly this dynamic conception of aesthetic preference formation. I regard this as a competitor to the Formation Thesis. Ivanova (2016, 2017 and 2020) adopted this approach. Following Poincaré, she provides an alternative explanation: understanding, which involves "an ability to grasp how the facts fit together" (Ivanova 2017, p. 6), is an epistemic value pursued by science, and "beauty is experienced when one has grasped how different and apparently disconnected phenomena are unified" (ibid., p. 6). The conceptual link between understanding and consistency should be evident. The advantage of this account lies in its ability to explain why values like simplicity and unity are regarded as stable "historical constants" unaffected by changing fashions (Montano 2014). However, it recalls Todd's objection: since understanding is an epistemic value, the satisfaction derived from it, even if expressed in aesthetic utterances, should naturally count as epistemic satisfaction. The understanding account thus gives us a reason to regard, for example, rigidity as an epistemic virtue. Skeptics could argue on this basis that rigidity is not genuinely an aesthetic property, and more broadly, that the alleged aesthetic preferences this account attempts to explain do not exist at all. Proponents of the understanding account might attempt to argue either that there is no substantive difference between aesthetic and epistemic satisfaction, or that understanding generates aesthetic satisfaction alongside epistemic satisfaction. In any case, they can at least respond that the understanding account is hermeneutic and focuses on scientists' aesthetic utterances, so they don't need to claim that these utterances are based on genuine aesthetic evaluations.

My assessment is that there is currently no strong reason to regard the Formation Thesis, introduced at the start of this section, as superior to its competitors as Ivanova's understanding account.

As a summary of Sections 5 and 6, (in)consistency can be understood as an aesthetic property—at least in the hermeneutic sense—because it answers the question, "Which feature of the theory elicits the scientist's aesthetic response?" This account aligns with scientists' own descriptions of theoretical beauty in terms of rigidity and naturalness. If, as conjectured, inconsistency provokes aesthetic displeasure and mistrust only when it occurs between a theory's form and content, then our meta-induction indeed treats (in)consistency as an aesthetic property. Put differently, in this meta-induction the link between (in)consistency and empirical success is mediated by the aesthetic value of harmony between form and content. Todd's concern—that scientists' aesthetic utterances may not reflect genuine aesthetic reactions—can be allayed by distinguishing subjective from objective disinterestedness and by drawing on neuroaesthetic findings. Although this matter ultimately becomes an empirical question, scientists' own attitudes give us reason for optimism. Finally, how aesthetic preferences for consistency actually form remains an open question, and there is no compelling reason to favor McAllister's meta-induction account over competing views such as Ivanova's understanding account.

7. Objections and Replies

Local-Consistency Objection: Consistency can be classified into local consistency and global consistency. A theory is globally consistent if it, as a set of propositions, does not include a proposition along with its negation. A theory is locally consistent if there exists an appropriate division of context, such that the theory remains consistent within each subcontext. As Brown (1990) shows, a theory may be locally consistent without being globally consistent. Thus, even if global consistency is not an epistemic requirement, local consistency might still be one.

That's right. I have to justify that local consistency is neither an epistemic requirement. Luckily, there are reasons to believe this. If Brown's (1990) analysis of Bohr's model is correct, the usability of Bohr's model depends, indeed, on its local consistency. However, this contextualization-based analysis faces significant challenges. Da Costa & French (2003, Chapter 5) argue that Brown (1990) overlooked the genuine "central inconsistency" of Bohr's model, which involves "the assertion that [an electron in the ground state] would not radiate energy and spiral into the nucleus as determined by classical physics." Taking this into consideration, they conclude that "Bohr's theory cannot be so easily broken down into distinct [subcontexts] to each of which one can systematically assign principles held as true" (p. 91). Moreover, in Planck's derivation of the law of black-body radiation, such a contextualization strategy was not employed (Norton 1987). Therefore, there is no strong evidence suggesting that the old quantum theories were indeed locally consistent.

Note that not all paraconsistent logics adopt this kind of contextualization strategy (see, for example, Da Costa *et al.* 1998). Thus, the failure of the contextualization strategy does not imply the failure of paraconsistent control. In sum, consistency—whether local or global—may be an epistemic virtue, but it is not an epistemic requirement.

Semantic Objection: How can we know if two *statements*, such as "Schrödinger's cat is alive" and "Schrödinger's cat is dead" are contradictory? The answer should depend on whether they can be both true : if they can, they are certainly not contradictory; if not, they maybe contradictory. Thus, semantically speaking, contradictions do not occur in nature, for once two "contradictory" statements are both confirmed, it ceases to be a contradiction. Therefore, the belief that contradictory theories cannot be true does not require meta-inductive justification.

My response is that this approach encounters difficulties when describing and interpreting actual scientific practice. Consider again the example of nonlinear quantum mechanics: on which ground Weinberg abandoned this theory? According to the semantic objection, Weinberg rejected it, for he *directly* judges that nonlinear quantum mechanics and relativity theory cannot both be true. Yet it remains unclear how this direct judgment itself can be grounded. Moreover, such a view faces a phenomenological objection: scientists undoubtedly believe that two theories cannot both be true precisely because they perceive them as contradictory, not the other way around.

The opponent's underlying concern might be that whether two statements constitute a contradiction depends on the chosen formalization scheme. Indeed, this is true. For already formalized theories, such as physical theories, this issue poses no significant problem. However, in fields with lower degrees of formalization—such as biology—this concern may indeed deserve greater attention.

Practice Objection: It is the practice, rather than the belief, that should be defended, even if the belief actually guides the practice. As an analogy, mathematicians may seek new axioms to solve a problem, such as proving the Continuum Hypothesis (which has been proven to be independent of ZF) and this practice can be supported by Platonism: if mathematical theories study mathematical objects which exist independently just as physical theories study physical objects, then this practice, like seeking new physical laws, can be fully justified. In fact, some mathematicians do hold Platonist beliefs. However, the truly important question remains whether the practice of seeking new axioms can be justified, and defending Platonism is merely one possible strategy, among others. We may justify this practice through other strategies, such as emphasizing that doing so enriches mathematics as an arsenal of empirical sciences. Similarly, what truly needs to be justified is merely the practice of physicists in discarding inconsistent theories. These practices can be justified by other reasons (such as consistent theories being simpler and more comprehensible).

In response, it is not clear to what extent scientists' rejection of theories which turn out to be inconsistent can be explained by the fact that consistent theories are more comprehensible. In any case, being less comprehensible does not by itself warrant abandonment. Even if the practice of discarding inconsistent theories can be defended on other grounds, whether the belief "inconsistent theories cannot be true" can itself be justified remains a question worth exploring—especially given that physicists, such as Weinberg, often act on the basis of this belief. Out of respect for scientific practice, examining this belief should be considered a priority.

Perhaps the objector is truly concerned with providing an explanation (rather than a justification) for existing scientific practices. They may say even if this belief (inconsistent theories cannot be true) does not hold, scientific practice could still be as it is today, conversely, even if this belief is true, scientific practice might not be guided by it and therefore, the truth value of this belief may not be important for the actual form of scientific practice. To explain why scientific practice is as it is today, we need to look elsewhere.

In response, whether this idea that inconsistent theories cannot be true is justified remains an important question for explaining scientific practice, because it is unclear whether scientific practice would be as it is today if this idea did not hold, as this idea has, in fact, been directly tested. According to Heisenberg's recollection, in the early days of the quantum era, Bohr doubted whether the inconsistencies in his model were ultimately inevitable:

> [B]ecause [Bohr] was so much impressed by these paradoxes which were apparently unavoidable, he counted always on the possibility, "Well, these paradoxes may even, in the long run, mean some kind of inconsistency which cannot be avoided."(Heisenberg in AHQP February 25, 1963)

If this belief failed such tests, then scientific practice would not be as it is today; conversely, the fact that it has passed such tests constitutes a justification for it.

8. Conclusion

The belief that beauty leads to truth is prevalent among contemporary physicists. Far from being a private faith, it operates as a methodological guiding principle, essentially in the context of justification. However, it is unclear how beauty should be understood here for this belief to be justified not merely as useful but as true.

In this article, I propose an interpretation of "beauty leads to truth" as "ugliness leads to falsehood," where "ugliness" refers to a lack of formal harmony, namely, a lack of consistency; in other words, "beauty leads to truth" is interpreted as "inconsistent theories cannot be true." As this article has shown, this conviction that inconsistent theories cannot be true is indeed utilized as a methodological principle in scientific practice.

Nevertheless, justifications could not be easily found, since this conviction is not a logical requirement, as paraconsistent logic allows us to work with inconsistent theories; indeed, scientists do use contradictory theories. Nor can it be justified by the absence of contradictions in direct observation, as formal inconsistency does not necessarily lead to observable contradictions; moreover, certain phenomena, such as quantum superposition, are indeed regarded by some scholars as describable using contradictory statements. It is also challenging to justify this belief based on the acceptance of a final theory or theory of everything, as the consistency of such a theory lacks independent support. It seems that the only non-circular justification for the conviction is based on a meta-induction : in the history of science, inconsistent theories have not been as successful as their consistent counterparts.

This meta-induction qualifies as an aesthetic induction, insofar as it concerns (in)consistency as an aesthetic property. My justification for this claim rests firstly on the idea that (in)consistency is an aesthetic property in a hermeneutic sense. Indeed, (in)consistency serves as the foundation for an account of theoretical beauty that rests on two main claims: (1) consistency, understood as a relation between a theory’s form (its constitutive representation) and its content (its derived representation), admits of degrees, with inconsistency spanning a range of lower consistency levels; and (2) the beauty of a theory depends (partly) on its degree of consistency. Several aesthetic properties frequently discussed in the physics literature—such as rigidity and (un)naturalness—can be accounted for within this framework. Moreover, under this account, inconsistency is perceived as ugly—and thus serves as an indicator of falsehood—only when it arises from a conflict between a theory's constitutive and derived representations. If a theory's constitutive representation is deliberately designed to be inconsistent, it may not be perceived as ugly, nor immediately judged as false. If this holds, then the link between (in)consistency and empirical success is mediated by an aesthetic value—namely, the harmony between form and content. What (in)consistency does, then, is aesthetic work, not logical work.

These so-called "aesthetic properties" and "aesthetic values" may ultimately prove to be epistemic properties and values in disguise. However, in the absence of empirical evidence supporting such skepticism—and given the frequent use of aesthetic vocabulary in the physics community, as well as scientists' own characterization of these features as aesthetic—we have good reason to remain optimistic about their genuinely aesthetic status.

Finally, how scientists—especially physicists—come to develop their aesthetic preference for consistency remains an open question.

Many other questions remain to be explored. First, it is unclear whether the first three justifications in Section 4 can be effectively refined. Furthermore, the historical justification relies on the claim that the successors of inconsistent theories have generally been more successful than the inconsistent theories themselves. However, what exactly does this "success" mean, and how should it be measured? To what extent can this success be attributed to consistency? To what extent does the history of physics support this claim? These questions are all well worth exploring.

More broadly, if the general statement "inconsistent theories cannot be true"—and some weaker version of it—ultimately proves difficult to justify, might we then have an argument in favor of dialetheism, according to which "there are true contradictions" (Priest, 1987, p. 4)? Conversely, if the statement holds, does it support a form of Pythagoreanism, according to which "the world instantiates a beautiful, harmonious pattern" (Bangu 2006, p. 408), and does it, to some extent, explain the "unreasonable effectiveness of mathematics in the natural sciences" (Wigner 1960)?

Additionally, the interpretation I present here is just one possible interpretation of "beauty leads to truth." Are there other plausible interpretations? And, if so, do these interpretations bear any connection to one another?

Finally, perhaps the greatest moral of this case is that some truths about nature reside less in any single theory and more in the comparison between theories, less within science itself and more within the history of science. What other truths of this nature might there be?

Acknowledgments

I warmly thank my supervisor, Marco Panza, and my friends Mia Joskowicz, Philippe Debroise, Xiangcheng Lu, Zhongrui Ying, and Sorin Bangu for their support and insightful discussions. I am grateful to the two anonymous reviewers for their detailed and constructive feedback. Thanks also to IHPST for hosting the Colloquium TTSM, and to the Chinese Scholarship Council for funding my research.

Bibliography

Achinstein, P. (2018). *Speculation: Within and about science*. Oxford University Press.

Arruda, A. I. (1980). A survey of paraconsistent logic. In A. I. Arruda, R. Chuaqui, & N. C. A. da Costa (Eds.), *Mathematical Logic in Latin America* (pp. 1–42). North-Holland.

Bangu, S. (2006). Pythagorean heuristics in physics. *Perspectives on Science, 14*(4), 387–416.

Barbieri, R., & Giudice, G. F. (1988). Upper bounds on supersymmetric particle masses. *Nuclear Physics B, 306*(1), 63–76.

Brown, B. (1990). How to be realistic about inconsistency in science. *Studies in History and Philosophy of Science, 21*, 281–294.

Brown, B. (1993). Old quantum theory: A paraconsistent approach. In *PSA 1992, Vol. 2*, Philosophy of Science Association, 397–411.

Brown, L. M., & Rechenberg, H. (1987). Paul Dirac and Werner Heisenberg—A partnership in science. In Kursunoglu & Wigner (Eds.), *Physics Today*, 23(4), 29–31.

Bueno, O. (2005). Dirac and the dispensability of mathematics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *36*(3), 465-490.

Bueno, O., & French, S. (2018). *Applying mathematics: Immersion, inference, interpretation*. Oxford University Press.

Currie, A. (2023a). Epistemic engagement, aesthetic value, and scientific practice. *The British Journal for the Philosophy of Science*, *74*(2), 313-334.

Currie, A. (2023b). Towards a new aesthetics of science: Aesthetic cultures and the processes and objects of regard. In M. Ivanova & A. Murphy (Eds.), The aesthetics of scientific experiments. Routledge.

Cutting, J. E. (2003). Gustave Caillebotte, French impressionism, and mere exposure. *Psychonomic Bulletin & Review*, 10, 319-343.

Da Costa, N. C. A., & De Ronde, C. (2013). The paraconsistent logic of quantum superpositions. *Foundations of Physics, 43*, 845–858.

Da Costa, N. C. A., & French, S. (2003). *Science and partial truth: A unitary approach to models and scientific reasoning*. Oxford University Press.

Da Costa, N. C. A., & Marconi, D. (1987). An overview of paraconsistent logic in the 80s. *Monografias da Sociedade Paranaense de Matemática* (Reprinted in *Logica Nova*, Akademie-Verlag, 1988).

Da Costa, N. C. A., Bueno, O. A. B., & French, S. (1998). The logic of pragmatic truth. *Journal of Philosophical Logic, 27*(6), 603–620.

Dedekind, R. (1930–32). *Gesammelte mathematische Werke* (Vols. 1–3). R. Fricke, E. Noether, & Ö. Ore (Eds.). Braunschweig: Vieweg; reprinted New York: Chelsea Publishing Company, 1969.

Dirac, P. A. M. (1970). Can equations of motion be used in high-energy physics? *Physics Today, 23*(4), 29–31.

Einstein, A. (1999). *Autobiographical notes*. La Salle, IL: Open Court.

Feynman, R. P. (1985). *QED: The strange theory of light and matter*. Princeton University Press.

Friedman, M. (2001). *Dynamics of Reason: The 1999 Kant Lectures at Stanford University.* CSLI Publications.

Heisenberg, W. (1963, February 25). [Oral history interview of Werner Heisenberg by Thomas Kuhn]. *Archive for the History of Quantum Physics*, Harvard University, Cambridge, MA.

Heisenberg, W. (1971). *Physics and Beyond: Encounters and Conversations.* Translated by Arnold J. Pomerans. New York: Harper and Row.

Hilgers, T. (2016). *Aesthetic disinterestedness: Art, experience, and the self.* Routledge.

Hossenfelder, S. (2018), Lost in Math: How Beauty Lead Physicists Astray. Basic Books.

Ivanova, M. (2016, online). Poincaré's aesthetics of science. *Synthese*. https://doi.org/10.1007/s11229-016-1069-1

Ivanova, M. (2017). Aesthetic values in science. *Philosophy Compass*, *12*(10), e12433.

Ivanova, M. (2020). Beauty, truth and understanding. In *The aesthetics of science* (pp. 86-103). Routledge.

Jaśkowski, S. (1969). Propositional calculus for contradictory deductive systems (communicated at the meeting of March 19, 1948). *Studia Logica: An International Journal for Symbolic Logic, 24*, 143–160.

Kragh, H. (1990). *Dirac: A scientific biography*. Cambridge University Press.

Kuipers, T. (2002), 'Beauty, a Road to the Truth', Synthese, 131: 291– 328.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science, 48*, 19–48.

Leibniz, G. W. (1989). *Philosophical essays* (R. Ariew & D. Garber, Trans.). Hackett Publishing Company.

McAllister, J. W. (1999). *Beauty and revolution in science*. Cornell University Press.

McDonnell, J. (2017). *The Pythagorean world: Why mathematics is unreasonably effective in physics*. Springer.

Menger, P.-M., & Verschueren, P. (Eds.). (2023). *Le Monde des mathématiques*. Seuil.

Meskin, A., Phelan, M., Moore, M., & Kieran, M. (2013). Mere exposure to bad art. *British Journal of Aesthetics*, 53(2), 139-164.

Montano, U. (2014). *Explaining beauty in mathematics: An aesthetic theory of mathematics* (Synthese Library, Vol. 370). Springer.

Murphy, A. (2023). Form and content: a defense of aesthetic value in science. Philosophy of Science, 90(3), 669-685.

Norton, J. D. (1987). The logical inconsistency of the old quantum theory of black body radiation. *Philosophy of Science, 54*, 327–350.

Norton, J. D. (1987). The logical inconsistency of the old quantum theory of black body radiation. *Philosophy of Science, 54*(3), 327–350.

Norton, J. D. (1993). A paradox in Newtonian gravitation theory. In *PSA 1992, Vol. 2*, Philosophy of Science Association, 412–420.

Penrose, R. (1974). The role of aesthetics in pure and applied mathematical research. *Bulletin of the Institute of Mathematics and Its Applications, 10*, 266–271.

Pincock, C. (2012). *Mathematics and Scientific Representation*. New York: Oxford University

Poincaré, H. (1913). *The foundations of science: Science and hypothesis, the value of* *science, science and method*. University Press of America, 1982.

Poincaré, H. (1970) [1905]. *La valeur de la science*. Paris: Flammarion.

Popper, K. R. (1940). What is dialectic? *Mind, 49*(196), 403–426.
Press.

Priest, G. (1987). *In contradiction*. Martinus Nijhoff.

Sartwell, C. (2024). Beauty. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2024 Edition). https://plato.stanford.edu/archives/fall2024/entries/beauty/

Schurz, G. (2008). The meta-inductivist's winning strategy in the prediction game: A new approach to Hume's problem. *Philosophy of Science, 75*(3), 278–305.

Smith, J. (1988a). Scientific reasoning or damage control: Alternative proposals for reasoning with inconsistent representations of the world. In *PSA 1988, Vol. 1*, Philosophy of Science Association, 241–248.

Smith, J. (1988b). Inconsistency and scientific reasoning. *Studies in History and Philosophy of Science, 19*, 429–445.

Steiner, M. (1989). The application of mathematics to natural science. *The Journal of Philosophy, 86*(9), 449–480.

Steiner, M. (1995). The applicabilities of mathematics. *Philosophia Mathematica, 3*(2), 129–156.

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

Todd, C. S. (2008). Unmasking the truth beneath the beauty: Why the supposed aesthetic judgements made in science may not be aesthetic at all. *International Studies in the Philosophy of Science*, *22*(1), 61-79.

Turner, D. D. (2019). *Paleoaesthetics and the Practice of Paleontology.* Cambridge University Press.

Von Neumann, J. (2018). *Mathematical foundations of quantum mechanics: New edition*. Princeton university press.

Weinberg, S. (2011). *Dreams of a final theory: The scientist's search for the ultimate laws of nature*. Vintage.

Wigner, E. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications of Pure and Applied Mathematics, 13*, 1–14.

Wilczek, F. (2005). Nobel Lecture: Asymptotic freedom: From paradox to paradigm. *Reviews of* Modern Physics, 77(3), 857-870.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of personality and social psychology*, 9(2p2), 1.

Zeki, S., Romaya, J. P., Benincasa, D. M. T., & Atiyah, M. F. (2014). The experience of mathematical beauty and its neural correlates. *Frontiers in Human Neuroscience, 8*, Article 68.