



# Phased Theory Construction and Reintegration of Evidence

## Decoupling Formal Structure from Interpretation (Clarity First)

Ground-up theory redesigns often begin by **stripping away all interpretive baggage** and focusing on a minimal formal framework. The goal is to establish a **self-consistent set of axioms, equations, or rules** in the “weakest possible sense,” without committing to physical meaning 1 2. For example, a project may postulate entities (domains, quantities, changes) with **no assumed geometry, time, or dynamics**, ensuring that **no hidden physical assumptions** creep in 3 2. This axiomatic approach prioritizes **internal consistency and coherence** of the theoretical structure above all. In fact, philosophers of science note that **internal consistency (no logical contradictions)** is a **fundamental criterion**, often **ranked even above empirical accuracy** when evaluating theories 4. A new framework must first *make sense on its own terms* – every element defined clearly and no internal conflicts – before worrying about matching external facts.

**Historical precedent:** *David Hilbert's axiomatization program* illustrates this decoupling. Hilbert approached physics like geometry – **starting from first principles and logic** – whereas Einstein simultaneously drew on physical intuition. “Einstein and Hilbert were engaged in qualitatively different enterprises... Hilbert’s goals were as much logical and epistemological... as they were physical,” a study explains 5. Hilbert formulated General Relativity via an elegant variational principle and attempted to derive electromagnetism from gravity within a tight axiomatic system. Though some physical assumptions (like Gustav Mie’s electromagnetic matter model) later proved wrong 6 7, Hilbert’s phased approach demonstrates the **first phase: get the formal structure right, ensuring clarity and logical consistency, before attaching physical interpretation**. Modern “first-principles” projects – from attempts at a **Theory of Everything** to novel field theories – echo this. They often **declare no physical meaning beyond the math itself** 8, aiming to **build a consistent mathematical machine** that could later be endowed with physical context. This clarity-first phase prevents us from inadvertently baking in old assumptions; it keeps the theory malleable and logically sound.

## Avoiding Premature Anchoring and Bias

A key reason to delay hypotheses and analogies is to **avoid premature anchoring** – the cognitive trap of seizing a familiar idea too early and unconsciously forcing the framework to fit it. In scientific work, **confirmation bias** can lead researchers to tweak a nascent model to recover expected outcomes, rather than allowing truly novel structures to emerge. To combat this, scientists borrow tactics from experimental methodology. For instance, “**blind analysis**” is used in particle physics and social science to prevent bias: analysts **hide the true outcomes or data labels** until their methods are finalized 9 10. This way, *no one knows if their interim result “looks right”*, so they focus on getting the logic and calculations correct. As Nobel laureate Saul Perlmutter notes, without blinding, researchers tend to scrutinize their work only when a result is surprising, often **nudging conclusions toward prior expectations** 11. By contrast, **in a blinded workflow the team debugs and cross-checks every analysis step without seeing the “real” answer**, ensuring **internal validity** independent of desired outcome 9 10.

Applied to theory construction, this means **temporarily bracketing out empirical feedback** – much like solving a puzzle without peeking at the answer. Theorists may choose **not to compare their emerging framework against familiar data or paradigms too soon**, to avoid subconsciously molding it to fit legacy theories. This *intentional ignorance* in early stages encourages exploring unconventional formulations that a too-early reality check might have unfairly discarded. It fosters **creativity and “conceptual wanderlust”**, as using analogies too early can inadvertently constrain thinking <sup>12</sup>. In practice, teams might establish an “**idea quarantine**”: initial conjectures and metaphors are documented but set aside while the formalism grows. By maintaining this discipline, one ensures that **the groundwork is influenced by first principles and logical necessity, not by forcing it to align with preconceived models**. In short, **stay fluid and skeptical of your initial hunches** – *treat them as possibilities, not constraints*, until the theory’s backbone is solid. This avoids the risk of locking into a suboptimal path due to early attachments.

## Epistemic Milestones Before Reintroducing Evidence

Even while empirical validation is on hold, a rigorous theory-building process will set **internal milestones** to decide when the framework is mature enough to face reality. These milestones act as **gates** between the exploratory phase and the reintegration phase. Common criteria include:

- **Logical Consistency Check:** All definitions and axioms should coexist without contradiction. For example, if adding a new axiom or equation causes a conflict with earlier ones that cannot be resolved, the theory isn’t ready. *Passing this check is critical* – a theory must meet this bar before any empirical use, as a theory with internal contradictions is essentially void. In practice, theorists often circulate a “state-of-the-theory” document ensuring each component is self-consistent <sup>13</sup> <sup>14</sup>. Only once consistency is verified (sometimes through formal proofs or computational checks) do they move on.
- **Complete but Minimal Structure:** The framework should cover all foundational aspects intended in this phase (e.g. definitions of space, time, fields, interactions) **in a minimal form**. The idea is to **finish the skeleton**: no gaps in the formal structure, but also no superfluous elements added merely to chase phenomena. For instance, one milestone might be writing down a Lagrangian or set of equations that encapsulate all assumed symmetries and principles, *without yet trying to parametrize empirical values*. Clarity on “what the theory says” must be achieved; any concept introduced later must not violate the established base <sup>15</sup> <sup>16</sup>.
- **Reproducibility of Known Limits or Identities:** Before integrating new evidence, the nascent theory can be tested against *logical or mathematical benchmarks*. Does it reproduce well-established *internal* constraints or special cases? For example, a quantum theory might be checked to ensure it reduces to classical mechanics in a certain limit (an external fact, but one expected of any viable theory), or a fluid-dynamics model might be tested on conservation-law identities. These are not full empirical tests, but **consistency with known theoretical limits**. Such checks serve as *proxy milestones* indicating the structure is robust enough to likely handle real-world data. As one paper on theory evaluation notes, scientists value **external consistency with related theories** as a virtue when direct evidence is sparse <sup>17</sup> <sup>4</sup>. If your framework can embed or mimic established theories in appropriate limits, it’s a sign that integrating actual evidence will be smoother.
- **Formal Derivation of Consequences:** Another milestone is reaching the point where the theory can *generate predictions or outputs on its own*, even if abstract. For example, deriving the dispersion

relation of your model equations <sup>18</sup> <sup>19</sup>, or proving a theorem about its behavior. If these derivations hold, they indicate the theory is internally fertile. More importantly, they create hooks for later comparison with experiment. A theory that cannot yet produce any calculable consequence (even a hypothetical one) is probably not ready – it's like a code that hasn't compiled yet. One **epistemic indicator of maturity** is when theorists can say: "If this framework is right, then *in principle* it entails X." Generating such entailments – without immediately checking them against nature – is a sign the theory has solidified into a testable form, ripe for empirical confrontation.

Meeting milestones like these builds confidence that **the idea is coherent and "idea-complete" enough** to justify bringing back the real-world context. Essentially, the theory graduates from a sandbox environment into a candidate for reality. At this transition, the "**empirical interface**" of the theory comes into focus – i.e. a clear mapping where theoretical quantities relate to observable quantities can now be defined (Guest (2024) calls this the empirical interface of a theory <sup>20</sup>). Crossing this point means the theory's creators have a well-defined way to translate formal elements into experimental or observational terms, at least in principle. Only when such translation exists (even if approximate) is reintegration feasible.

## Timing the Reintegration of Hypotheses and Evidence

**When** should prior evidence, intuitions, and conjectures be folded back into the picture? The sweet spot is when the **core framework is stable but still adaptable** – after the above milestones are reached, but *before* the theory ossifies or over-specializes. At this juncture, one can gradually reintroduce external information in a controlled way. Best practices for this reintegration phase include:

- **Incremental Empirical Alignment:** Rather than a flood of data, introduce one empirical element at a time and see how the theory handles it. For example, plug in a well-known constant or effect (say, Newton's gravitational law or a specific particle property) and check for consistency. Successful historical cases often did this piecemeal. *Albert Einstein's development of General Relativity* is a classic example: only after formulating generally covariant field equations did he **fold in known anomalies** like Mercury's perihelion shift to validate the framework. The theory, by design, had the mathematical capacity to produce a perihelion advance; Einstein checked and found it matched the observed 43" per century – a huge empirical triumph, but *one tested after the theory's structure was set*. From there he felt confident to predict new evidence (light bending), fully reintegrating the empirical reality with his theoretical edifice. The lesson is to **pick pivotal pieces of evidence as tests or guides** once you think your theory can accommodate them, rather than tuning the theory to every data point from the start.
- **Use Prior Conjectures as Hypothesis Tests, Not Axioms:** Early conjectures or analogies that inspired the project can now be revisited – but critically, treat them as hypotheses *to be tested* within the new framework, not as additional axioms. For instance, if a theory was developed in isolation but was inspired by an analogy to neural networks, the reintegration phase might involve checking if the theory indeed exhibits a network-like behavior under certain conditions. If yes, great – the analogy holds *as a result*, not by construction. If not, it might indicate either the analogy was misleading or the theory needs refinement; in either case, one avoids having *forced* the analogy in prematurely. At this stage, prior ideas serve as **benchmarks or heuristic guides**, helping to identify what empirical phenomena or intuitions the theory *should* eventually explain. Researchers might maintain a **checklist of legacy conjectures** ("does it recover classical thermodynamics? does it mirror this symmetry?") and start ticking them off as the theory's predictions solidify.

- **Scheduling feedback loops:** Many teams formalize the reintegration step as part of their workflow. For example, in **computational modeling, verification and validation** are distinct phases: first verify the model internally (no bugs, solves equations right), then validate against real data <sup>21</sup> <sup>22</sup>. A similar template can be used in theory-building. **Plan multiple integration points** – perhaps a small one early on as a sanity check and a bigger one later. One practical template is a “**theory sprint**” followed by a “**reality checkpoint**”. In the sprint, you develop the formalism freely for a set period or until hitting predefined milestones. Then comes a checkpoint: pause and compare any derived results with known facts or experimental findings. Importantly, define in advance what will be compared and what criteria count as “close enough” versus a breakdown. This guards against confirmation bias – you won’t cherry-pick what to compare after the fact, but stick to a planned test. If the theory passes the test (or can be adjusted without destroying internal consistency), that evidence is now integrated, and you proceed to the next sprint. If it fails badly, you revisit the formal structure. **Templates like this impose discipline** on reintegration: evidence is neither ignored indefinitely nor allowed to constantly reshape the theory ad-hoc; it enters at scheduled, well-considered junctures.
- **Empirical Risk and Reward Assessment:** Deciding *when* to incorporate evidence also involves judging the **risks of waiting vs acting**. If a particular piece of evidence could outright falsify a core aspect of the theory, some argue it’s better to test it sooner rather than later (to avoid wasted effort on a dead-end). Others counter that if the theory is still plastic, a premature falsification might be inconclusive – you could be testing a still-imperfect version. A balanced strategy is to **classify evidence by its criticality**. “High-criticality” tests (those that would make you abandon or drastically revise the theory) might be scheduled only after you have high confidence in the structure – essentially as final hurdles. “Low-criticality” evidence (more peripheral phenomena) can be integrated earlier to guide tuning. In foundational physics, for example, a theory might postpone confronting something like the value of the cosmological constant (high stakes) until late, but might early on ensure it can reproduce simpler well-known limits like inverse-square laws or Lorentz invariance (lower stakes guiding features). This triage ensures you **fold in guidance from prior knowledge** in a sensible order – building up from easy-to-fit phenomena to the most potentially refuting tests. Each successful integration increases confidence that the theory is ready for the next, more stringent exam.

In essence, **reintroducing hypotheses and evidence is done gradually, strategically, and on the theory’s terms**. By the end of the reintegration phase, the theory that started “in a vacuum” of pure thought is fully back in touch with reality – ideally explaining known facts in a clearer way and pointing to new predictions.

## Balancing the Risks: Under- vs Over-Integration

Striking the right moment (and manner) to reintegrate prior intuitions is a delicate balancing act. Both **premature and belated integration carry risks**:

- **Dangers of Over-Integrating Too Early:** If you anchor the theory to empirical notions too soon, you risk **stifling its development** and possibly baking in **false assumptions**. History provides stark examples. The *Ptolemaic epicycle model* of the solar system grew extremely convoluted by adjusting circles-upon-circles to match every new astronomical observation. By over-integrating each new data point into the model immediately, the Greeks and medieval astronomers locked themselves into a

complex geocentric system that “worked” for predictions but obscured the simpler heliocentric truth. Premature commitment to the idea that Earth was the center made them add epicycles rather than question the core assumption. This illustrates how **anchoring on a hypothesis (Earth-centered cosmos) from the outset** led to a cumbersome framework. Only when Copernicus and Kepler were willing to **break the early anchor** did a coherent, simpler theory emerge. In modern times, theoretical physics has arguably seen a form of this with certain fashionable ideas. An observer of the field, Sabine Hossenfelder, notes that for decades physicists became *over-committed* to concepts like supersymmetry or “naturalness” – and kept **shifting their predictions** in response to null experimental results, rather than reconsidering the initial premise <sup>23</sup>. This *moving of goalposts* – e.g. saying “*maybe supersymmetry appears at a higher energy we haven’t reached yet*” each time the current collider finds nothing – is a symptom of early anchoring. The risk is a **degenerating research program** that adjusts itself to avoid refutation indefinitely, losing credibility. In short, integrating evidence *too early* (or rigidly) can make a theory brittle or lead to self-deception, where one forces the theory to fit known facts at the expense of elegance or honesty. It can also squander the theory’s potential to surprise us with genuinely new predictions (since it was molded to only produce expected results).

- **Dangers of Under-Integrating (Delaying Too Long):** On the flip side, postponing empirical confrontation for too long can lead a project into an **esoteric cul-de-sac**. A theory left in splendid isolation might grow ever more internally elaborate but with no guarantee of relevance. The obvious example is **superstring theory** and related “Theory of Everything” attempts. After initial successes in ensuring mathematical consistency (e.g. anomaly cancellation that requires extra dimensions), string theory developed for decades with rich internal structures (branes, dualities, multiverses) **before making any contact with experiment**. To date it still lacks definitive testable predictions. This illustrates the danger of a *perpetual first-phase* – one can keep refining the mathematics and generating new postulates (many “compactifications” or exotic scenarios) without ever looping back to check: is any of this real? The result can be **stagnation**. Hossenfelder calls the current situation in fundamental physics “stagnation” not because of lack of ideas, but because so many ideas have flourished disconnected from experimental feedback and thus “**nothing is moving**” toward solid empirical confirmation <sup>24</sup> <sup>25</sup>. Another cautionary tale is *Arthur Eddington’s later work*: he attempted a grand unified theory in the 1930s-40s by pure reasoning (assigning profound meaning to mathematical coincidences), largely ignoring the new quantum evidence. He became increasingly isolated and his theory never aligned with experiment, effectively dying with him. **Delaying integration too long risks building an internally beautiful but irrelevant castle in the sky**. It may also cause one to miss the window to adapt the theory—small discrepancies that could have been fixed early might compound into huge problems. Moreover, external credibility wanes: a framework that produces no testable outcomes for a long time can lose scientific interest (critics might label it “not even wrong”).

The key is to navigate between Scylla and Charybdis: **neither freeze the theory around old ideas prematurely, nor let it drift indefinitely unmoored from reality**. Successful frameworks often oscillate: a period of free exploration and **divergent thinking**, then a checkpoint to **converge with evidence**, then another creative phase, and so on. Sir Karl Popper described science as *conjectures and refutations* – one must have the courage to conjecture freely, but also the courage to test and potentially refute those conjectures. **Avoid emotional attachment** to early hypotheses: treat a beloved analogy as expendable if Nature disagrees. Conversely, if evidence is lacking or ambiguous, don’t force-fit it; sometimes it’s wise to put it aside and refine the theory further until a clearer test arises. The reintegration phase is not one

moment but an ongoing negotiation: a healthy theory **continuously incorporates new data** as it becomes available, yet remains **grounded in its core principles** so it doesn't shapeshift arbitrarily.

## Historical and Cross-Disciplinary Insights

Many scientific endeavours have effectively managed phased theory construction:

- **Einstein's General Relativity (1915):** *Phase 1:* from 1907–1915 Einstein built the theory guided by principles (equivalence principle, general covariance) and mathematical consistency, at times **suppressing immediate empirical alignment**. Notably, he guessed forms of field equations by internal criteria (covariance, energy-momentum conservation) and only after finalizing them did he check known anomalies (Mercury's orbit) – which the theory nailed. *Phase 2:* post-1915, he boldly *predicted* new phenomena (gravitational light bending, redshift) as reintegration of evidence, rather than using them to shape the theory initially. This phased approach (first ensure a consistent structure, then confront observations) paid off handsomely. Interestingly, Einstein had earlier made a “premature integration” misstep: in 1913 he and Grossmann produced an incomplete theory (the “Entwurf” theory) partly because Einstein was then convinced general covariance might not hold in reality (anchoring to a flawed hypothesis about coordinate conditions). It was only by letting go of that anchor that he arrived at the correct theory in 1915. So even within Einstein's process we see the need to avoid locking in an early assumption and to trust the math once it proved elegant and consistent.
- **Hilbert vs. Einstein – Two Routes to Relativity:** As discussed, Hilbert's axiomatic route built the theory *from the top down* with formal rigor, while Einstein's route was more physically intuitive. **Both succeeded in different ways**, but Hilbert's experience also highlights reintegration issues. He had essentially the same gravitational field equations as Einstein. However, he attempted to also derive electromagnetism within the same axiomatic system (folding in a unification hypothesis too early) and used an outdated electromagnetic model in doing so <sup>6</sup>. The result was that part of his theory was a “hopelessly failed physics” <sup>7</sup>. In modern terms, Hilbert over-integrated a conjecture (that EM is just gravitation in disguise via Mie's theory) into his otherwise brilliant formalism. The success part – his independent discovery of Einstein's equations – shows the power of formal first principles. The less successful part underscores that *when* and *how* to integrate additional hypotheses (like unification schemes) is crucial.
- **Paul Dirac's Approach to Theory-Building:** Dirac was legendary for **prioritizing mathematical beauty and internal elegance over immediate empirical fit**. He famously stated, “*It is more important to have beauty in one's equations than to have them fit experiment... if one has really a sound insight, one is on a sure line of progress. If there is not complete agreement with experiment, one should not allow oneself to be too discouraged, because the discrepancy may be due to minor features that will get cleared up with further development of the theory.*” <sup>26</sup>. This philosophy led him to predict the existence of antimatter from the negative-energy solutions of his relativistic wave equation at a time when no such particle was known. He trusted the theory's internal logic and symmetry, and experimental evidence (the positron) was found later, vindicating him. Dirac's case illustrates a **successful delayed integration**: he didn't force the theory to exclude negative energies (which would have made it fit known particles only); instead, he let the theory be and expected evidence to catch up. His **epistemic milestone** was the beauty and logical tightness of the equation – once that was in place, he was willing to bet on nature conforming eventually. Dirac's strategy should be used

with caution (not every “beautiful” equation finds an experimental home), but it exemplifies the potential reward of *under*-integrating just long enough. The criterion here was a strong one: the theory was so constrained and elegant that Dirac had confidence in it despite an empirical gap. When a theory has that kind of inevitability (in his case, uniting quantum mechanics and special relativity forced the equation form), it might be correct to **trust the theory until evidence catches up** – a calculated risk.

- **Computational Physics & Model Workflows:** Outside fundamental physics, **multi-stage workflows are standard in computational science**. In complex simulations (climate models, astrophysical codes, etc.), scientists separate **verification (internal consistency) and validation (empirical accuracy)**. They first ensure the model correctly solves the intended equations (e.g. by checking conservation laws, grid convergence, reproducing known analytic solutions) – analogous to building a theory for internal coherence. Only then do they compare the model output to real-world data and observations to tune parameters or assess validity <sup>21</sup> <sup>22</sup>. This separation is explicit in fields like aerospace engineering (where design teams might not test a prototype in the real world until simulations and component tests are coherent) and in algorithm development (where one might prove an algorithm’s correctness before running it on real datasets). The lesson for theoretical frameworks is clear: **do your “code review” and unit tests on the theory first** – check all the math, edge cases, and self-consistency – then do the “wind tunnel tests” of comparing with experiment. If a mismatch occurs at that point, you can be more confident it’s the theory’s conceptual issue, not a trivial error, much as a failed validation in simulation indicates a deficiency in the model assumptions (not a mere coding bug). This staged approach is essentially a concrete template for reintegration.
- **Wolfram Physics Project (2020–present):** Stephen Wolfram’s recent attempt to find fundamental physics via hypergraph rewrite rules is a contemporary example of phased theory construction. In the first year of the project, the team predominantly “**explor[ed] the very rich structure of [the] models and their connections to existing theoretical frameworks,**” while explicitly stating they were only *on the path* to making direct experimental predictions <sup>27</sup>. They treated their evolving model (a new class of discrete space-time rules) as a formal system to be understood deeply in its own right – mapping how it yields general relativity and quantum mechanics in the appropriate limits, for instance – **before** diving into detailed empirical tests. Only once the internal behavior was charted did they begin discussing potential observable consequences (like specific cosmological or particle signatures). This mirrors a **two-phase workflow**: first build and “**climb the mountain**” of **formal complexity to see the view**, then later look outward to reality. The Wolfram project also adopted an *open notebook* approach, documenting every misstep and correction <sup>28</sup>. This transparency actually enforces phased rigor: it’s harder to quietly fudge the theory to fit a fact if the entire development is being logged in public. Their process underscores the value of **new tools** (computational exploration, automated verification of properties) in executing a phased strategy. By extensively using computer algebra and search to explore consequences of their rules, they accelerated the internal-consistency phase. The lesson here is that **modern computational methods can assist phased theory-building** – e.g. using automated theorem provers or large-scale simulations to test internal consistency and derivable phenomena *before* matching to real data.
- **Cross-Disciplinary Design Thinking:** In fields like engineering, design, and even organizational theory, a parallel concept exists: **divergent vs. convergent thinking stages**. In design thinking, one first goes broad – generating ideas without constraint (divergent, analogous to the hypothesis-free formal exploration) – and later converges by applying filters like feasibility and requirements

(analogous to reintroducing evidence and practical constraints). This staged creativity can be seen in any major project: for example, early in the SpaceX rocket program, engineers toyed with many designs (some quite radical) in simulation and theory; only after narrowing to a promising architecture did they start integrating extensive test data from engine firings and flights to refine the design. The general principle is widely applicable: **separate the creative generation of solutions from the critical evaluation of solutions in time**. By not trying to do both at once, you avoid killing off good ideas too early or, conversely, avoid charging ahead with an unvetted idea. Successful innovation often cycles between these modes. Science at the frontier is no different – it requires the imaginative formulation of theories *and* the cold-eyed testing against nature, but attempting both simultaneously is often counterproductive. Instead, **allocate a phase for free invention and a phase for merciless criticism** (with possible mini-cycles in between).

## Conclusion

Ground-up redesign of a theoretical framework benefits from a **phased methodology**. In the initial phase, **purity of thought and internal coherence reign supreme** – the theory is nurtured in isolation, like a young plant under careful conditions, to ensure it grows strong roots (first principles, logical consistency). Extraneous hypotheses, analogies to prior systems, and even known facts are held at arm's length to prevent bias and premature fixation. As the structure matures, reaching key **epistemic milestones** of consistency and completeness, the process enters a new phase: **reintegration**. Here the theory is progressively exposed to the wild – prior conjectures are revisited, empirical data is introduced – first gently, then with increasing rigor. Throughout, the theorist must balance open-minded creativity with critical restraint, knowing when to **suspend judgment** and when to **insist on alignment with reality**.

Historical successes show that this balance is achievable. The reward for patience and careful timing is a framework that is **clear and internally sound**, yet not unmoored from the world it seeks to describe. By avoiding the twin perils of **anchoring too early** (which can entrench biases and superficial fixes) and **delaying too long** (which can lead to elegant irrelevance), one guides the theory's evolution optimally. Practically, this means enacting policies in the research workflow – from blind analyses to scheduled validation checkpoints – that enforce the separation of concerns, followed by their reunion at the proper time. The outcome is a phased construction where **clarity and consistency are achieved first, providing a solid foundation upon which evidence and meaning can be layered**. Such a strategy not only yields a more robust final theory, it also provides a clearer view of *why* the theory works, since its internal rationale was developed free from ad-hoc adjustments. In sum, the best practices across disciplines counsel us to **build the house before painting it**: get the architecture right (even if in bare form), then gradually add the colors of reality. By the end of the process, one hopes to have a structure that is both **internally beautiful and externally true** – a theory that stands firmly on first principles and yet *embraces the world* it intends to explain.

**Sources:** Clear methodological insights and examples of phased theory development have been discussed in various contexts [2](#) [4](#) [9](#) [26](#) [27](#) [23](#), highlighting the importance of internal consistency first and strategic reintegration of empirical elements.

---

[1](#) [2](#) [3](#) [13](#) [14](#) [15](#) [16](#) skeleton.docx

file:///file\_0000000067f4722f916970eedcf4c948

4 17 20 Lessons for Theory from Scientific Domains Where Evidence is Sparse or Indirect | Computational Brain & Behavior

<https://link.springer.com/article/10.1007/s42113-024-00214-8>

5 6 7 Hilbert's 'Foundations of Physics': Gravitation and electromagnetism within the axiomatic method - ScienceDirect

<https://www.sciencedirect.com/science/article/abs/pii/S1355219807000780>

8 18 19 crft\_whitepaper.pdf

[file:///file\\_000000000a3871f5859f8ab0cc0311c6](file:///file_000000000a3871f5859f8ab0cc0311c6)

9 10 11 'Blind analysis' could reduce bias in social science research - Berkeley News

<https://news.berkeley.edu/2015/10/08/blind-analysis-could-reduce-bias-in-social-science-research/>

12 Conceptual wanderlust: How to develop creative supply chain ...

<https://onlinelibrary.wiley.com/doi/10.1111/jscm.12305>

21 22 Verification and validation - Wikipedia

[https://en.wikipedia.org/wiki/Verification\\_and\\_validation](https://en.wikipedia.org/wiki/Verification_and_validation)

23 24 25 Sabine Hossenfelder: Backreaction: The present phase of stagnation in the foundations of physics is not normal

<http://backreaction.blogspot.com/2018/11/the-present-phase-of-stagnation-in.html>

26 Paul A. M. Dirac Quotes on Equation from - 45 Science Quotes - Dictionary of Science Quotations and Scientist Quotes

[https://todayinsci.com/D/Dirac\\_Paul/DiracPaul-Equation-Quotations.htm](https://todayinsci.com/D/Dirac_Paul/DiracPaul-Equation-Quotations.htm)

27 28 The Wolfram Physics Project: A One-Year Update—Stephen Wolfram Writings

<https://writings.stephenwolfram.com/2021/04/the-wolfram-physics-project-a-one-year-update/>