Scientific Realism and Convergent Evidence

Roland Rosmont

Scientific realists commonly rely on the No Miracles Argument (NMA): the idea that the best explanation for scientific success is that scientific theories are at least approximately true (Putnam, 1975; Psillos, 1999). This argument has a strong intuitive appeal. If mature scientific theories do not describe an objective reality, it would be hard to explain how they, for instance, can predict new phenomena or enable successful technological intervention. To reject this, one has to accept that science's track record is nothing more than a miraculous coincidence — a position that few, if any, seem prepared to defend.

Still, NMA, in its standard formulation, faces three long-standing criticisms. First, its central notion of scientific success is too broad. From a philosophical perspective, predictive accuracy, technological control, empirical fit, and interpretive coherence all support realism in different ways and degrees. The second criticism comes from Laudan's pessimistic induction (Laudan, 1981). The history of science is filled with examples of theories that persisted for extended periods before being discarded. This demonstrates that success is no assurance of truth. Furthermore, Stanford argues that scientists often overlook alternative explanations when developing or evaluating their theories (Stanford, 2006). In other words, simply ruling out only the currently conceived alternatives cannot justify confidence in present theories.

These objections do not disprove realism. What they do show, however, is that the NMA needs a clearer definition of success. The problem, then, is that NMA does not distinguish clearly between successes that support realism and successes that do not. Until there is a clearly articulated understanding of "success," it will not be possible to conclude that the success that requires an explanation actually supports realism.

This essay argues that reconstructing NMA strengthens scientific realism if the explanandum — what counts as success — is defined explicitly. To do so, we need a critical mass of evidence that is reconstructible, methodologically diverse, and open to critical evaluation. That evidence should be gathered under conditions set up to expose shared sources of error. The framework for this approach is *Reconstructible Empirical Realism* (RER), which holds that scientific realism is justified by convergent evidential support.

Why Scientific Success Is Not Enough

The classical version of NMA asserts that the approximate truth of scientific theories explains their successes. In its unqualified form, three main issues challenge this conclusion.

The problem of empirical success. Van Fraassen's constructive empiricism identifies one principal limitation. A theory can succeed in achieving empirical adequacy by fitting recorded data across a variety of contexts, but it cannot justify its claims about unobservable entities (van Fraassen, 1980). Thus, empirical success and realist warrant are not logically dependent. A theory may accurately predict phenomena, lead to technological innovation, and organise collected data into reliable patterns, while its central theoretical assumptions remain unjustified. NMA does not actually resolve this gap.

The pessimistic induction. Laudan (1981) argues that an overly simplistic inference from success to truth is based on a partial understanding of the history of science. Many historically successful theories, such as the caloric theory of heat and phlogiston chemistry, were eventually revised or abandoned. Neither retains its central theoretical assumptions within its successor sciences. The list, says Laudan, "could be extended *ad nauseam*" (1981, p. 33). Perhaps the most sophisticated realist response — Psillos' "*divide et impera* move" (1999, p. 108) — is a significant partial solution, in which realist commitment is restricted to those working propositions that have genuinely contributed to the success of a theory. But drawing a clear and non-*ad hoc* distinction between the working and idle parts of a theory has proven difficult (Lyons, 2006).

Unconceived alternatives. Stanford (2006) takes Laudan's list of challenges a step further. He contends that the problem is not simply that scientists choose incorrectly among available theories — the problem of underdetermination — but that they typically fail to conceive of the alternatives from which better-confirmed successors will eventually be drawn. Stanford calls this "transient underdetermination" (Stanford, 2006, p. 17) and shows that it occurs even when a theory has successfully passed numerous tests against all currently imagined alternatives. Reconstructibility of evidential chains can expose the reasoning from observation to theory, but it cannot reveal theoretical possibilities that have not yet been realised.

These objections show that the concept of success is under-analysed and does not offer a solid foundation for realism. What NMA needs is a clearer account of the specific kind of success that favours a realist interpretation. The task, therefore, is to identify the evidential conditions under which scientific success acquires realist significance. RER is proposed as an account of those conditions.

Reconstructibility and Empirical Content

Between a theoretical statement *S* and its observation lies an empirical chain of procedures. These intermediate steps may include model development, experimental design, measurement and calibration, data management, and statistical analysis. The organisation and recording of these procedures determine the evidential strength of the result (Franklin, 2013; van Fraassen, 1980). Some steps are robust, while others are not, yet all are philosophically significant.

A chain of evidence supporting claim *S* is considered reconstructible when each step is documented well enough for qualified researchers to assess its reliability and determine the impact of different analytical choices. Good critical tracing relies on reproducible research practices, including preregistered protocols, openly shared data, and reports of unsuccessful or null findings (Ioannidis, 2005; Nosek *et al*., 2015). The recent replicability crisis illustrates the importance of such documentation (Fanelli, 2018). Several findings, once accepted, lost credibility following independent review. Empirical performance alone cannot fully establish the validity of scientific claims.

RER distinguishes degrees of reconstructibility by asking how completely a proposition can be traced through documented empirical practice. *Empirically empty* claims lack solid, reconstructible chains of evidence. Such claims rely solely on anecdotal or speculative links to experience. A *weakly reconstructible* claim has some empirical support, but one or more important steps in the inferential sequence remain unclear or only partly documented. By contrast, a *strongly reconstructible* claim is supported by a well-documented evidential route that skilled investigators can review, test, and, when necessary, reproduce. This typology replaces the traditional empiricist dichotomy of meaningful and meaningless propositions with a graded account of evidential depth. It also establishes the primary philosophical distinction between empirical significance and realist warrant, developed below.

The ability to reconstruct evidence supports its empirical significance. But this, by itself, does not establish realism. Thus, two epistemic levels need to be kept separate: a *content level*, concerning whether, and how firmly, a claim is grounded in inspectable practice, and a *warrant level*, concerning whether approximate truth is epistemically justified as the best explanation of that grounding. To move from level one to level two, we need strong reconstructibility, but that alone does not suffice. Even a well-documented experiment can still contain systematic mistakes, and a research tradition can preserve the same background errors across successive studies. As a result, realism requires further structure beyond reconstructibility, supplied by the conditions specified in the next section.

RER as a Stronger No Miracles Argument

RER reconstructs NMA by giving its *explanandum* a more precise form. In the traditional version, approximate truth is presented as the best explanation of a theory's success. In RER, the target is narrower and more demanding: approximate truth explains why several independent but fallible evidential routes continue to converge on the same claim after researchers have tested for shared sources of error and alternative explanations.

For this realist inference to be warranted, a minimum of three conditions must be met.

To be reconstructible, evidential chains must be theoretically specified, systematically verified, and exposed to independent reanalysis. Sensitivity tests using alternative analytical strategies and

preregistered reviews strengthen the evidential chain by limiting *post hoc* flexibility. Without these practices, even impressive published results provide only weak support for realism.

Method coherence adds a further requirement. Support for a claim should arise from evidential routes that use different instruments, assumptions, experiments, and statistical methods. Such variation exposes the claim to different risks of failure. Partial independence, in this context, is practical rather than metaphysical. Evidential chains that rely on different calibration strategies, modelling choices, or collaborative structures are less likely to share the same source of error.

Wimsatt emphasises the epistemic value of robustness: when there are "independent alternative ways of deriving a result, the result is always surer than its weakest derivation" (Wimsatt, 1981, p. 132). Multiple distinct causal mechanisms make it harder to attribute agreement to a single undiscovered flaw. Convergence, therefore, reduces the plausibility of shared-error explanations and gives the evidence epistemic force. Although empirical adequacy can explain fit within a particular domain, persistent steadiness across independently structured investigations is more readily explained by the theory's approximate truth.

Alternative exclusion requires researchers to examine whether convergence results from calibration artefacts, publication bias, or shared theoretical assumptions. The failure of such explanations strengthens the realist argument. The relevant phenomenon is the persistent convergence of independent, fallible procedures on the same result.

A Bayesian critique of the NMA argues that it collapses into the "base-rate fallacy" by neglecting prior probabilities (Howson, 2000, p. 54). This objection applies most directly to the classical formulation. RER places the explanandum within a more specific context. The relevant question is how probable a convergence pattern would be if the theory were not approximately true. As alternative explanations are identified, tested, and found inadequate, the likelihood of the observed convergence under anti-realist hypotheses decreases relative to the realist hypothesis. Given appropriate priors, this shifts the posterior likelihood in favour of realism (Sprenger, 2016).

Van Fraassen's remaining challenge concerns empirical adequacy. A theory can match observations in a given domain without being approximately true. Yet empirical adequacy has difficulty explaining why independently generated evidential routes continue to coincide under conditions designed to reveal shared error. Approximate truth provides a fuller explanation, since it accounts for both the observational fit within each route and the stability of that fit across intentionally varied investigative paths. This broader consideration allows the realist hypothesis to explain the evidential pattern better than a purely instrumentalist interpretation (van Fraassen, 1980; Psillos, 1999).

The Higgs Boson as a Paradigm Case

The discovery of the Higgs boson demonstrates RER in practice. The evidence for its existence developed over several years of theoretical prediction, experimental investigation, and data analysis (Franklin, 2013). The case for the Higgs boson rests on a stable pattern of results from various lines of inquiry.

Strong reconstructibility was achieved through detailed documentation, published calibration records, and extensive independent reanalysis. ATLAS and CMS used different detector geometries, electronics systems, and algorithms, while the principal decay routes were subject to distinct systematic fluctuations. Even so, both experiments observed a signal near 125 GeV with a combined significance of 5 sigma, corresponding to a probability of less than one in a million that the excess resulted from statistical variations alone. This convergence, therefore, appeared across partially independent evidential routes exposed to different risks of failure. Blind analyses, control regions, and statistical testing were used to rule out known sources of noise. Statistical fluctuation, calibration artefacts, and analysis bias were all considered, yet none accounted for the overall pattern of results.

The stability of the signal across procedures designed to detect artefactual agreement is what calls for explanation. In this case, approximate truth, understood as the existence of a Higgs-like scalar field, provides the strongest account of the evidential situation. A sceptic may still challenge aspects of quantum theory or raise underdetermination concerns about the interpretation of the Standard Model. These objections are philosophically significant, but they do not remove the distinctive epistemic standing of the Higgs claim. Its warrant rests on an unusually demanding evidential process of the kind RER requires.

String theory, by comparison, occupies a different epistemic position. Its central claims are currently without empirical chains that are comparable in reconstructibility, methodological coherence, and a record of excluding alternatives (Ellis & Silk, 2014; Dawid, 2013). It is still a heuristically valuable research programme, though its central claims do not presently warrant realist commitment in the procedural sense specified by RER.

Replies to Major Objections

Pessimistic induction. *RER changes the reference class rather than dissolving the objection.*

Many of Laudan's paradigm cases were supported by narrow, methodologically homogeneous evidence and were largely insensitive to the kind of scrutiny of alternative explanations that RER demands. That said, some historical theories — Fresnel's wave optics and certain formulations of ether mechanics — did exhibit partial methodological coherence and non-trivial evidential convergence. RER increases the explanatory burden on anti-realist accounts while remaining fallibilist about scientific theories. Where a

theory meets all three RER conditions, the anti-realist must explain its eventual revision and why independent, fallible, and systematically scrutinised procedures converged so stably before that revision occurred. That burden is more philosophically demanding than the one imposed by bare predictive success. RER is explicitly fallibilist: its conditions do not immunise theories against overthrow, but they make the pessimistic induction a less tractable instrument against current science, since the evidential structures it would need to indict are more demanding than those it historically targeted.

Unconceived alternatives.*RER addressespractical underdeterminationwhile acknowledgesthat logical underdetermination is unavoidable*.

A genuine rival within the RER framework must be integrated into reconstructible empirical chains of comparable depth, exhibit methodological coherence equal to that of the competing account, and withstand equally stringent tests of alternative explanations. Speculative alternatives that have not reached such procedural maturity do not carry the same epistemic weight, and the asymmetry in procedural depth is itself evidentially significant. RER does not promise immunity from future unconceived rivals; it shows that the current realist commitment rests on a more demanding and more stable evidential base than traditional NMA presupposes (Stanford, 2006). The possibility of future rivals remains an abstract epistemic consideration, but absent procedurally mature alternatives, it does not undercut the comparative rationality of current realist commitment.

Selective and structural realism. *RER does not defend realism about whole theories*.

RER defends a realist view of the components that actually sustain the reconstructible and convergent evidential chains. In this respect, it is structurally aligned with Psillos' *divide et impera* strategy, which directs realist commitment at the working propositions, causal mechanisms, and theoretical statements that fuel success-generating mechanisms (Psillos, 1999). It is also compatible with Worrall's structural realism: what RER identifies as warranting commitment is the evidentially supported relational structure confirmed across independent chains, rather than any particular characterisation of the entities' intrinsic nature (Worrall, 1989). Consequently, RER reaches selective and structural realism as a *procedural* conclusion, not as an *ad hoc* rescue from the pessimistic induction but as the natural epistemic output of its three conditions applied to mature scientific cases.

Conclusion

RER retains the core intuition of NMA but sharpens the explanandum, controls the inference process, and grounds the realist claim in a more demanding evidential structure.

Success in science does not, by itself, justify realism. An engaging scientific observation, a well-cited prediction, or an institutionally celebrated result may all fail to meet RER's conditions. Any realist who

turns to such cases weakens their position. The relevant phenomenon is more demanding: multiple independent evidence chains must converge while competing explanations fall short. When this occurs, the pattern itself requires explanation.

Approximate truth is not the only possible explanation. However, it offers the strongest account of why theory and evidence align, and why that alignment remains stable across studies. That is the explanatory advantage of realism over empirical adequacy.

RER does not defend realism based on scientific success in general. It defends realism only where rigorous inquiry repeatedly converges despite various methods, tools, and theoretical perspectives. And when such a pattern exists, then realism remains the most epistemically defensible response to a specific and demanding evidential situation.

Bibliography

Dawid, R. (2013). *String Theory and the Scientific Method*. Cambridge University Press.

Ellis, G., & Silk, J. (2014). Scientific Method: defend the integrity of physics. *Nature (London)*, 516(7531), pp. 321–323.

Fanelli, D. (2018). Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences - PNAS*, 115(11), pp. 2628–2631.

Franklin, A. (2013). *Shifting Standards: Experiments in Particle Physics in the Twentieth Century*. University of Pittsburgh Press.

Howson, C. (2000). *Hume's Problem: Induction and the Justification of Belief*. Clarendon.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science*, 48(1), pp. 19–49.

Lyons, T. D. (2006). Scientific realism and the stratagema de *divide et impera*. *The British Journal for the Philosophy of Science*, 57(3), pp. 537–560.

Nosek, B. A., Alter, G., Banks, G. C., *et al*. (2015). Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science (American Association for the Advancement of Science)*, 348(6242), pp. 1422–1425.

Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. Routledge.

Putnam, H. (1975). *Philosophical papers. Vol. 1, Mathematics, Matter and Method*. Cambridge University Press.

Sprenger, J. (2016). The probabilistic no miracles argument. *European Journal for Philosophy of Science*, 6(2), pp. 173–189.

Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press.

Van Fraassen, B. C. (1980). *The Scientific Image*. Clarendon.

Wimsatt, W.C. (1981). Robustness, reliability and overdetermination. In Campbell, D. T., Brewer, M. B., & Collins, B. E. (Eds.) *Scientific Inquiry and the Social Sciences* (pp. 124–163). Jossey-Bass Publishers.

Worrall, J. (1989). Structural realism: the best of both worlds? *Dialectica*, 43(1–2), pp. 99–124.