

Meta-Empirical Confirmation: Addressing Three Points of Criticism

Richard Dawid

February 3, 2020

Abstract

I respond to three points of criticism that have been raised against the concept of meta-empirical confirmation.

1 Introduction

It has been argued in Dawid (2013), Dawid (2006, 2007, 2009, 2016, 2018, 2019) and Dawid et al. (2015) that three arguments of *non-empirical confirmation* were capable of significantly increasing trust in a scientific theory's viability in the absence of empirical confirmation. In the present discussion note, I want to respond to points of criticism presented by Smolin (2014), Chall (2018) and Menon (2019), which are directed specifically at individual arguments of non-empirical confirmation. I will not discuss other, more general lines of criticism of the approach (some of which I have engaged with elsewhere) as expressed for example in Ellis and Silk (2014), Ellis (2017), Cabrera (2018), Dardashti (2019), Hossenfelder (2019), Oriti (2019), and Rovelli (2019). Nor will I be able to discuss all points made in the three papers addressed. I will focus on one core point in each case, which in my understanding represents the most important issue raised in the given text.

2 Zooming in on Three Arguments of Meta-Empirical Confirmation

The concept of non-empirical theory assessment denotes lines of reasoning that aim to generate a significant degree of trust in a theory's viability in the absence of empirical confirmation. While Dawid (2013) singles out a small set

of specific arguments that can be particularly powerful tools of non-empirical confirmation, it does not rule out the existence of other forms of significant non-empirical theory assessment. Recently, a wider range of such strategies has been addressed in the literature (see e.g. Peebles (forthcoming)). In this light, the present paper will use the more specific term "meta-empirical confirmation" to denote specifically the three arguments of non-empirical confirmation discussed in Dawid (2013) and closely related arguments. I define arguments of meta-empirical confirmation as arguments that increase the trust in a theory's viability by inferring limitations to scientific underdetermination from observations about the way the scientific research process has played out. Those observations are about the world but don't amount to empirical evidence for the theory because they are not of the kind that can be predicted by the theory in question.

Scientists deploy three specific arguments of meta-empirical confirmation when evaluating their theories in the absence of sufficient empirical confirmation: i) The no alternatives argument (NAA): Scientists tend to trust a theory if they observe that, despite considerable efforts, no alternative theory that can account for the corresponding empirical regime is forthcoming. ii) The unexpected explanation argument (UEA): Scientists tend to trust a theory if they observe that the theory turns out to be capable of explaining significantly more than what it was built to explain. iii) The meta-inductive argument (MIA): Scientists tend to have increased trust in a theory that fulfills the first or the first two criteria if it is their understanding that previous theories in their research field that satisfied those criteria had usually turned out empirically successful once tested.

As indicated above, these arguments are based on assessing the spectrum of possible alternatives to the theory in question. In each of the arguments, the meta-level observation about the research process serves as an indicator that there is a scarcity of possible alternatives to the given theory. If a scientist has plausible reasons to infer from observations about the research process that possible conceptual alternatives to a known theory are probably very scarce or absent, this provides an epistemic basis for trusting that theory.

3 First criticism: Non-empirical confirmation is too easy to achieve

Lee Smolin (2014) argues that meta-empirical confirmation is so flexible that scientists with a sufficient degree of ingenuity can always use it to support the theory they prefer. As an example, Smolin aims to demonstrate that

loop quantum gravity (LQG) can be confirmed meta-empirically just as well as string theory. He claims that all three arguments of meta-empirical theory confirmation can be deployed also for confirming LQG.

Smolin views LQG and string theory as rival approaches that can both be supported by a no-alternatives argument if one adds specific requirements: full unification in the case of string theory and manifest background independence in the case of LQG. While the former requirement is not met by LQG, the latter is not met by a perturbative approach to string theory.

In order to develop his argument, Smolin needs to make two problematic steps, however. First, he shifts the level of analysis from the level of theories to the level of research programs. Second, he assumes that methodological preference can provide a basis for NAA.

It is fair to call LQG and string theory rivalling research programs that choose different strategies for addressing the problem of quantum gravity. LQG focuses on the quantization of gravity while string physics approaches the question based on a universal theory of all interactions. However, research programs don't get confirmed. What can get confirmed is an individual theory. As it stands, LQG and string theory do not constitute theories about the same class of phenomena. String theory covers all interactions while LQG does not.¹ String theory therefore has the more extensive intended domain. It follows that, while string theory is an alternative to LQG as a theory of quantum gravity, LQG today does not constitute an alternative to string theory as a theory of all interactions. A no-alternatives argument therefore can be applied to string theory without referring to full unification in the narrow sense of relying on just one fundamental physical principle. All that is needed is the specification of the theory's intended domain: string theory is claimed to be the only known theory that covers all known interactions.

While string theory therefore can find support from a NAA based on specifying the theory's intended empirical domain, LQG can only find such support based on an additional requirement that disqualifies string theory.² Smolin suggests the requirement of manifest background independence. In order to understand whether introducing a requirement of this kind is ac-

¹LQG may some day in the future be developed into a theory that describes all interactions. If that can be done, the resulting theory would play out within the LQG research program but require additional posits and therefore constitute a different theory than the one investigated today. Such a theory, if it did not merge with string theory and covered the observed phenomenology of high energy physics in an explanatorily sufficiently satisfactory way, would amount to a rival theory to string theory.

²For the sake of the argument, we set aside the point that LQG is just one of a range of approaches of quantizing gravity in a manifestly background independent way that haven't been shown to be physically equivalent.

ceptable, we need to discuss an aspect of NAA that, though implicit in the way the argument has been presented in Dawid (2013), may not have been sufficiently emphasized.

An epistemically relevant no-alternatives claim needs to be based on a specified class of physical phenomena that allegedly cannot be represented by an alternative theory (such as the class of all known interactions in the case of string theory). Once such a class of phenomena has been specified, a no-alternative claim can be justified in the following way: Assuming that science works at all in the given context, there must be a scientific theory that can account for the given class of phenomena. Therefore, if only one scientific theory can account for the given class of phenomena, that theory must be viable.

Manifest background independence does not denote a class of phenomena. Rather, it constitutes a methodological choice. Any attempt to build a no-alternatives case based on a methodological choice immediately faces a regress problem, however: in order to assess the significance of such a NAA, one needs to assess the probability that other theories can successfully represent the phenomena of quantum gravity without making the choice of manifest background independence. Since there is no good way of assigning probabilities to methodological choices, the corresponding no alternatives argument deflates to a mere statement of methodological preference and provides no basis for assessing the relevant spectrum of unconceived alternatives. Therefore, contrary to Smolin's claim, the condition of manifest background independence cannot be used fruitfully in a NAA.

The lack of a genuine NAA for LQG also substantially reduces the basis for a MIA for the theory, since a MIA acquires strength only in conjunction with a no-alternatives case.

Smolin also discusses ways in which the argument of unexpected explanation (UEA) applies to LQG. There is no reason to deny that unexpected explanation can arise in LQG. To the extent it does, it amounts to non-empirical theory confirmation for LQG. Specific unexpected explanatory interconnections that arise for both string theory and LQG would only have reduced confirmation value for each theory, however, since their occurrence in different theoretical approaches would suggest that they are related to a deeper characteristic of theorizing about quantum gravity rather than to an individual theory.

Assessing the substance and significance of unexpected explanations provided by a specific theory, be it string theory or LQG, is a matter of careful analysis. This is not the place to carry out such an investigation. It is my understanding that the complex and far-reaching web of unexpected explanations encountered in the context of string theory is not matched by what

one finds in the case of LQG. But be this as it may, the physicist's careful assessment of the extent of unexpected explanations must provide the basis for assessing UEA's significance.

To conclude, NAA is not applicable to LQG, which in turn removes the basis for applying MIA. UEA might be applicable to some extent but it is doubtful whether it has comparable strength to the string theory case.

4 Second Criticism: Arguments of Non-empirical confirmation are structurally flawed.

Cristin Chall (2018) has argued for structural flaws of UEA and MIA. Chall starts his argument against UEA by claiming that "the UEA is essentially a non-empirical version of the no miracles argument" (Chall 2018, p132). On that basis, he argues that the significance of UEA is threatened by the possibility that unexpected explanatory interconnections could also be provided by a more fundamental theory than the one supposedly supported by UEA. Chall then asserts that Dawid (2013) counters the described threat in the case of string theory by recourse to a final theory claim: if the theory is final, no more fundamental theory exists. Based on this reconstruction of UEA, Chall argues against the argument's validity: UEA is deployed to establish the theory's viability; a final theory argument can only be established based on the assumption that the theory is viable; if UEA itself relies on a final theory claim, the line of reasoning is viciously circular. Moreover, if UEA is applicable only to theories that generate a final theory claim it is irrelevant for most of scientific reasoning.

With regard to MIA, Chall claims that the argument is not coherently applicable if the meta-inductive support from the success of other theories in the research field relies on theories to which the theory under scrutiny is a successor theory. Chall points out that the agreement of the predecessor theory's predictions with the data needs to be retained by any viable successor theory (in some low energy limit of that successor theory). Therefore, the agreement between that data and the predecessor theory always amounts to empirical confirmation (accommodation) of the successor theory. But if the agreement between the predecessor theory and the data has already been fully factored in as empirical confirmation, it cannot provide additional *meta-empirical* evidence for the successor theory's viability.

Given that string theory is a universal theory, any empirically viable theory in high energy physics deals with physics covered by string theory and

thus must be its predecessor theory in Chall's sense. Therefore, Chall argues, MIA is in principle inapplicable to string theory. Moreover, he claims that most good cases of comparable theories that could enter an MIA argument are predecessor theories, which renders MIA mostly inapplicable.

Chall's criticisms are based on an insufficient appreciation of two core elements of meta-empirical confirmation: (i) the importance of the distinction between local and global underdetermination, and (ii) the importance of the distinction between the role of experimental data at the ground level and at the meta-level.

Chall's statement that UEA is a non-empirical version of the no-miracles argument (NMA) ignores an important difference between the two arguments: they aim to support substantially different claims. NMA is deployed in support of scientific realism. In asserting a theory's approximate truth, scientific realism makes the global claim that the theory will *never* be replaced by a successor that contradicts the theory's core ontic commitments. UEA, to the contrary, is a local argument. It is deployed to support a theory's viability *within a given empirical horizon*. If the theory gets superseded at a higher energy scale that lies beyond the empirical horizon, the local viability claim supported by UEA remains intact. The possibility of a more fundamental theory whose characteristic empirical implications lie beyond the considered empirical horizon therefore does not threaten UEA. Thus no final theory claim is needed to block any such threat.

What does threaten UEA is the possibility of an underlying principle that is not specific to the given theory but applies to a wider group of possible theories about the given intended domain. If such an underlying principle exists, one of the alternatives may be viable, the theory under scrutiny may be false, and the unexpected explanatory interconnection may arise due to the underlying principle as instantiated in the viable alternative theory.

As pointed out in Dawid (2013), this threat cannot be countered within the context of UEA. It can be controlled, however, by deploying MIA and/or NAA in conjunction with UEA. Mutual reinforcement between the three arguments can in the end achieve significant confirmation.

Chall's argument thus goes wrong in two respects. His argument for the need of a final theory claim as a foundation for UEA is based on the erroneous understanding that UEA aims to support a global (realist) claim. And his analysis of the mechanism of UEA ignores the important point that the significance of UEA cannot be argued for without reference to one or both other kinds of meta-empirical confirmation.³

³One should note that final theory claims can further strengthen meta-empirical confirmation (see Dawid 2013, p153f). They are not necessary, however, for making

In order to see the problem with Chall’s argument against MIA, it is crucial to acknowledge the structural difference between ground level observations E and meta-level observations F. E denotes empirical data that happen to be in agreement with theory H ’s predictions. The corresponding F (in the MIA case) denotes the observation that H had been without known alternatives for a considerable time despite intense search for alternatives and ended up being empirically confirmed by data E.

Chall is right that the agreement between a predecessor theory’s predictions and data E needs to be retained by a successor theory. Therefore, once one has checked that the successor theory can reproduce its empirically successful predecessor in some low energy limit, E indeed must not be counted a second time as independent evidence for the successor theory. Meta-level evidence F, however, is not of a kind that can be retained by the successor theory. F represents a contingent fact about the way the research process played out with respect to the development and testing of the predecessor theory. That contingent fact is neither predicted by the predecessor nor by the successor theory. Therefore, F can have confirmation value for the successor theory beyond the observation that the predecessor theory is consistent with data E. Contrary to Chall’s claim, MIA thus can very well be applied to successor theories.

To conclude, neither of Chall’s two criticisms poses problems for meta-empirical confirmation once one adequately accounts for the full structure of the argument.

5 Third Criticism: the Gerrymandering Charge

Menon (2019) raises a specific issue related to NAA. Dawid et al. (2015) presents a proof that NAA amounts to confirmation in a Bayesian framework under very weak and plausible conditions. As pointed out in the paper, this proof in itself does not establish the significance of confirmation based on NAA. Confirmation could be marginal and therefore, though formally realized, unhelpful for scientific theory assessment. This situation is in agreement with the claim made in Dawid (2013) that modes of meta-empirical confirmation achieve significance only in conjunction.

While Menon does not dispute the results of Dawid et al. (2015), he raises the question as to how natural and plausible significant confirmation would be in the scenario it describes. He reaches the conclusion that significant

UEA+NAA+MIA significant and work in a different way than Chall suggests.

confirmation can only be achieved by NAA based on a deliberately chosen – “gerrymandered”, as he calls it – set of priors for the hypotheses Y_i (where Y_i states that there is a number i of possible scientific theories that can account for data E). Specifically, Menon’s analysis shows that the significance of NAA requires high priors for very low i -s that quickly fall off for higher i -s. This, Menon argues, is not a generic choice. Moreover, it is a choice that seems to be begging the question, since high probabilities for low numbers of alternatives is exactly what NAA is supposed to establish.

Menon’s formal analysis is sound and accurate. The distribution of priors he singles out is of the kind one indeed needs for significant NAA. What I will contest in the following is his verdict that this choice of priors is gerrymandered in scientific contexts where NAA is taken seriously by scientists.

The crux of the matter, once again, boils down to the fact that NAA requires support from other forms of meta-empirical confirmation to get off the ground. In the absence of such support, NAA indeed faces the problem Menon describes. Scientific reasoning in a mature scientific field, however, is distinguished by the fact that NAA *can* be supported by MIA. And, as we will shortly see, MIA in a predictively successful research field very plausibly generates the distribution of probabilities for Y_i -s that is demanded by Menon. These probabilities, which constitute posteriors with regard to updating under MIA, then can serve as priors for a NAA argument.

In order to see this point, let us first look at one individual case of predictive success of a theory H . Consider a theory H that has been developed based on a data set E_0 and successfully predicts data E . Let F_{nc}^H denote the observation that (i) H was developed without data E entering the construction process and (ii) H successfully predicts E . The question we want to ask is the following: if we start with a given spectrum of priors $P(Y_i^H)$, what are the posteriors $P(Y_i^H|F_{nc}^H)$?

When addressing the question, we individuate theories based on their distinct predictions within a given empirical horizon. E rules out all but one of the theories that can possibly be developed in agreement with E_0 and thereby establishes the empirical viability of theory H within the given empirical horizon. Moreover, we assume that developing a theory in accordance with the available data E_0 before experiments E were carried out resembles making a random pick from those theories that are consistent with E_0 . In other words, before measuring E we have no reason to assume that the predictively successful theory will be easier or less easy to find than any other theory that is consistent with E_0 .

Let us now start with the following fairly conservative choice of priors for Y_i^H :

$$P(Y_i^H) = 1 / \left(\sum_{j=1}^N j^{-1} \right) i \quad \forall \quad i \leq N \quad (1)$$

$$P(Y_i^H) = 0 \quad \forall \quad i > N. \quad (2)$$

For this choice of priors, $P(Y_i^H) \rightarrow 0$ for all i in the limit $N \rightarrow \infty$, but finite numbers of alternatives are not dogmatically excluded.⁴

We then have:

$$P(F_{nc}^H | Y_i^H) = 1/i \quad (3)$$

and

$$P(F_{nc}^H) = \sum_{i=1}^{\infty} (P(Y_i^H) P(F_{nc}^H | Y_i^H)) = \sum_{i=1}^{\infty} i^{-1} P(Y_i^H) \quad (4)$$

Equations (1), (3), (4) and Bayes' theorem lead to

$$P(Y_i^H | F_{nc}^H) = \frac{P(F_{nc}^H | Y_i^H) P(Y_i^H)}{P(F_{nc}^H)} \quad (5)$$

$$= (i^{-2}) \left(\sum_{j=1}^N j^{-1} \right)^{-1} \left(\left(\sum_{k=1}^N k^{-2} \right) \left(\sum_{j=1}^N j^{-1} \right)^{-1} \right)^{-1} \quad (6)$$

$$= (i^{-2}) \left(\sum_{j=1}^N j^{-2} \right)^{-1} \quad (7)$$

$$\lim_{N \rightarrow \infty} P(Y_i^H | F_{nc}^H) = \frac{1}{\zeta(2) i^2} \quad (8)$$

where ζ is the Riemann ζ -function. For the lowest i -s, we get:

$$\begin{aligned} \lim_{N \rightarrow \infty} P(Y_1 | F_{nc}^H) &\cong 0,61 \\ \lim_{N \rightarrow \infty} P(Y_2 | F_{nc}^H) &\cong 0,15 \\ \lim_{N \rightarrow \infty} P(Y_3 | F_{nc}^H) &\cong 0,067 \end{aligned} \quad (9)$$

⁴We will say a little more about the plausibility of this choice of priors later on.

which amounts to an even more pronounced preference of small i -s than Menon demands. On this basis, we can see already that the probability spectra for numbers of unconceived alternatives demanded by Menon are not entirely unnatural. Rather, they roughly resemble what one would plausibly assume for any empirically viable scientific theory that was developed without guidance by the data that eventually confirmed it.

Note, however, that what is stated in Equations (9) are the posterior probabilities for the Y_i -s of an individual theory H once its empirical viability has been established. What we are actually interested in are prior probabilities for a theory that has not yet been empirically tested. In order to get there, we need the meta-inductive step of MIA. This step relies on a series S of all known cases that share relevant characteristics K (such as not having any known alternatives before being empirically tested) and for which we know the theories' eventual predictive successes. A strong tendency of predictive success in S will suggest that satisfying conditions K is correlated with a spectrum of Y_i -s that is conducive to developing ("picking") predictively successful theories. This will lead to further Bayesian updating where the posteriors of empirically tested theories that satisfy K will influence the priors of new so far untested ones.

Without entering the intricacies of that process, it is very plausible to conjecture the following: Only substantial priors for small i -s allow for the expectation that the corresponding theory has a significant chance of predictive success. Therefore, if the members of S show a strong tendency of predictive success (even if there are some failures), this must substantially increase priors for low i -s for new as yet untested theories that satisfy K ^{5,6}

Each step towards a new theory that satisfies K raises the question, however, whether the new theory is indeed sufficiently similar in all relevant respects to the members of S to justify an inference from the the average Y_i spectrum we have inferred for members of S to the Y_i -s for the new theory. Doubts in this regard will significantly decrease credences for low i -s for the new empirically unconfirmed theory compared to the averaged spectrum of

⁵Note that this will, in turn, push the posterior credences for small i -s even beyond the values of Equations (9) for theories that satisfy K and have turned out to be predictively successful.

⁶We can now understand the plausibility of choosing Equation (1) as our default set of priors. Updating based on one instance of predictive success leads from an even distribution of priors $P(Y_i) = 1/N$ to the distribution of Equation (1). If we use the previous theory's posteriors as the priors for the next theory under scrutiny, even a long but finite series of failures will not move the probability distribution significantly away from Equation (1). This suggests that, in a scientific field where predictive success has occurred at all, it is reasonable to choose priors that favor small i -s at least to the extent indicated by Equations (1).

Y_i -probabilities for a member of S . The power of MIA, while considerable in cases where members of S have a strong tendency of predictive success, therefore must remain limited.

This is the reason why NAA can get off the ground based on a convincing case of MIA but nevertheless has a distinct and important role to play. The following overall scenario emerges. MIA can support the significance of NAA type reasoning in a research field by looking at comparable cases where empirical viability could eventually be checked. However, the significance of MIA is constrained by the fact that it is a generalization argument that relies on an inference class S of theories that are different in many respects to the theory we want to confirm. NAA can now step in and rely on investigating the strength of a no alternatives verdict specifically in the given case, carefully searching for alternatives, developing a better theoretical understanding of the obstacles to such attempts, etc. This individual view on the specific context under scrutiny can then, based on the set of priors extracted from MIA, lead to further significant confirmation of the given theory.

To conclude, we find that the Y_i distribution needed for a significant NAA doesn't have to be gerrymandered but can be naturally generated by MIA in a predictively successful research field. NAA is still necessary as an independent argument due to the limited effectiveness of MIA.

6 Conclusion

All three lines of criticism addressed in this note raise relevant and important issues. Answering them was based on pointing out that significant meta-empirical confirmation involves a number of conceptually non-trivial elements: the NAA needs to be based on the specification of the intended empirical domain rather than on methodological or conceptual requirements; all three arguments of meta-empirical confirmation rely on the careful distinction between local and global claims and on the specific status of meta-level observations; finally, the arguments can be significant only by working in conjunction. Whether and to what extent significant confirmation can be achieved in the end depends on the specific characteristics of the individual scientific case. I hope to have demonstrated, though, that none of the lines of criticism discussed threaten the core of meta-empirical confirmation.

Acknowledgements

I'm very grateful for very helpful remarks by Casey McCoy and Tushar Menon on draft versions of this paper. Funding for this research was provided by the Swedish Research Council (project number 1598801).

References

- Cabrera, F. (2018): "String Theory, Non-Empirical Theory Assessment and the Context of Pursuit", Synthese, online first.
- Chall, C. (2018): "Doubts for Dawid's Non-empirical Theory Assessment", *Studies in History and Philosophy of Modern Physics* 63:128-135.
- Dardashti, R. (2019): "Physics Without Experiments?". In Dardashti, R., R. Dawid and K. Thebault (eds) *Why Trust a Theory? -Epistemology of Fundamental Physics*, Cambridge University Press.
- Dawid, R. (2006): Underdetermination and Theory Succession from the Perspective of String Theory, *Philosophy of Science* 73/3, 298-322.
- Dawid, R. (2007): Scientific Realism in the Age of String Theory, *Physics and Philosophy* 11: 1-32.
- Dawid, R. (2009): On the Conflicting Assessments of the Current Status of String Theory, *Philosophy of Science*, 76/5, 984-996.
- Dawid, R. (2013): *String Theory and the Scientific Method*, Cambridge: Cambridge University Press.
- Dawid, R. (2016): Modelling Non-Empirical Confirmation, in Ippoliti, E. T. Nickles and F. Sterpetti (eds.), *Models and Inferences in Science*, pp 191-205, Springer 2016.
- Dawid, R. (2018): "Delimiting the Unconceived", *Foundations of Physics* 48(5):492-506.
- Dawid, R. (2019): "The Significance of Non-Empirical Confirmation in Fundamental Physics". In Dardashti, R., R. Dawid and K. Thebault (eds) *Why Trust a Theory? -Epistemology of Fundamental Physics*, Cambridge University Press.

- Dawid, R. S. Hartmann and Jan Sprenger (2015): "The No Alternatives Argument", *The British Journal for the Philosophy of Science* 66(1), 213-234.
- Ellis, G. (2017): "Physics on Edge", *Inference* 3:2.
- Ellis, G. and J. Silk (2014): "Defend the Integrity of Physics", *Nature* 516: 321-323.
- Hossenfelder, S. (2019): *Lost in Math: How Beauty leads Physics Astray*, Basic Books.
- Menon, T. (2019): "On the Viability of the No Alternatives Argument", *Studies in History and Philosophy of Science A*, 76:69-75.
- Oriti, D. (2019): "No Alternative to Proliferation". In Dardashti, R., R. Dawid and K. Thebault (eds) *Why Trust a Theory? -Epistemology of Fundamental Physics*, Cambridge University Press.
- Peebles, J. (forthcoming): *Cosmology's Century - An Inside History of our Modern Understanding of the Universe*, Princeton University Press.
- Rovelli, C. (2019): "The Dangers of Non-Empirical Confirmation". In Dardashti, R., R. Dawid and K. Thebault (eds) *Why Trust a Theory? -Epistemology of Fundamental Physics*, Cambridge University Press.
- Smolin, L. (2014): "Review of 'String Theory and the Scientific Method' ", *American Journal of Physics* 82:1105.