Rethinking Holism and Underdetermination*

Florian J. Boge† ‡

Abstract: Mature scientific hypotheses are confirmed by large amounts of independent evidence. How could anyone be an anti-realist under these conditions? A classic response appeals to confirmational holism and underdetermination, but it is unclear whether traditional arguments succeed. I offer a new line of argument: If holism is interpreted as saying that the confirmation of every part of a hypothesis depends on the confirmation of the whole hypothesis, we must formulate conditions under which the confirmation received by the whole can be transferred to its parts. However, underdetermination suggests that relevant conditions are typically not met. If this is true, the confirmation received by the whole remains bounded by the priors for the parts, and we lack compelling reasons to believe substantive hypotheses based on evidence beyond the degree to which the posits involved in them are antecedently believed. A rejoinder comes from selective realism: If some posit is preserved throughout theory change, it is confirmed beyond the degree to which the containing hypothesis is. However, the variant of holism considered here exactly implies that we cannot confirm such posits in isolation. As I will show, the realist is thus forced into a dilemma: Either she succumbs to the holistic challenge, or she must embrace meta-empirical facts, such as the posit's recurrence, as confirmatory.

Keywords: holism • underdetermination • selective realism • meta-empirical evidence

1 Introduction

Mature scientific hypotheses are confirmed by large amounts of independent evidence. How could anyone be an anti-realist under these conditions? A classic response to this question is based on confirmational holism (CH), the thesis that, in order to test a given hypothesis, we have to embed it into a whole network of hypotheses. And so only that network as a whole can ever be confirmed. But different, rival networks are thinkable that may all 'save the same phenomena'. Hence, the hypotheses in question cannot really receive any confirmation.

*This is the accepted manuscript version of a paper forthcoming in *Synthese*. The doi will be provided when available.

†Institute for Philosophy and Political Science, TU Dortmund University, Emil-Figge-Str. 50, 44227 Dortmund, Germany. florian-johannes.boge@tu-dortmund.de

‡Lamarr Institute for Artificial Intelligence and Machine Learning

The final conclusion I have attempted to draw from CH here is the key impact it is supposed to have on realism: If no given hypothesis can really be confirmed, then it is unclear why we should believe what it says. Crucially, the argument expressed in the preceding paragraph builds on a variant of the underdetermination problem (UP):[1] That evidence alone does not determine which theories we should entertain.

Both Duhem (1914) and Quine (1951) embraced a connection between UP and CH, but both were also rather "inexplicit" about the nature of this connection (Okasha, 2002, 307), and it is actually all but clear whether the argument of the first paragraph succeeds (see Okasha, 2002, 309–10). So why think that CH poses a serious challenge for realism at all, and what, if anything, is the connection to UP?

Using a probabilistic model of confirmation, I will show how CH can seriously challenge realism—and already at the 'local' level of a single hypothesis or theory. As I shall argue, sufficiently general and substantive scientific hypotheses[2] have logico-semantic parts – that is, non-trivially imply meaningful claims – which inevitably involve theoretical commitments that impact the level of confirmation the overall hypothesis can receive. To give an example: The Maxwell equations, which lie at the heart of our present understanding of electromagnetic phenomena, make claims not only about the relation between, but also the nature of, electric currents and electromagnetic fields. But following a standard approach to confirmation, it can be shown that the confirmation of the relations – which are crucial to the equations' success – depends on the confirmation of the claims about the natures, and vice versa. This is a clear statement of holism, but already at the local level of a single hypothesis or theory.

Combined with elementary facts about probabilistic confirmation, this means trouble, because it implies that we might end up in a situation where we can only confirm hypotheses to the extent that we *already believe* what they entail.[3] However, one might counter that, under certain circumstances, the increasing confirmation of the whole will also increase the degree of confirmation received by the parts. Call that 'holistic top-down

[1]Throughout, I will focus on transient underdetermination (Sklar, 1975), for radical underdetermination essentially amounts to skepticism (Laudan and Leplin, 1991, 450) and transient underdetermination has proven its worth in the realism-debate (see Stanford, 2023). The only tentative example of interesting, real-world radical underdetermination are the rival accounts of the quantum domain. They are tentative, because there might be empirical consequences to tease some of them apart (Das and Dürr, 2019). There might also be sufficient overlap between them (Fraser and Vickers, 2022) to allow realism about central theoretical claims (however, see also Boge, 2023, in this connection). Similarly, I will focus only on 'contrastive', rather than 'holistic', underdetermination (Stanford, 2023): The latter concerns the fact that if evidence is in contradiction with a prediction, we can blame any one out of the interwoven hypotheses used to devise the test. It thus challenges disconfirmation and with it, scientific rationality rather than realism (Stanford, 2023; Wray, 2018).

[2]I will sometimes omit the qualifier.

[3]I will show all this in detail later.

confirmation'. In this way, both whole and parts may be confirmed in concert, and beyond the degree of prior belief. While this is a plausible response, I shall argue that certain conditions are plausibly necessary for such holistic top-down confirmation to work, and that extant cases of underdetermination of theory by evidence suggest that we lack reason to think that all these necessary conditions are met in present science. Hence, by UP, the strategy fails.

The only serious response to this I see comes from selective forms of scientific realism (SSR): Given that a posit has figured in different, successively developed rival theories, it is confirmed beyond the degree of confirmation received by the given theory or hypothesis as a whole. However, insofar as local holism is correct, SSRists need to appeal to meta-empirical evidence in order to justify their response, as I will demonstrate. Hence, the realist faces a dilemma: Either she must succumb to the holistic challenge, or she needs to embrace meta-empirical evidence as confirmatory.

In short, this is my argument then: Holism is fleshed out to say that the parts of a hypothesis cannot be confirmed independently, but only via confirmation of the whole. However, for logico-mathematical reasons, no hypothesis can be confirmed beyond the degree of confirmation of its parts. Thus, if certain parts do not receive confirmation from the confirmation of the whole, holism limits the degree to which hypotheses can be confirmed *at all*. One might attempt to avoid this conclusion exactly by transferring confirmation down from the whole to its parts, so that both increase in credibility *in concert*. But underdetermination suggests that strategies for so transferring confirmation downwards, thereby also raising the degree to which hypotheses can be confirmed as a whole, typically fail.

The novelty of this argument lies in the fact that we now see a rigorous trajectory from holism via underdetermination towards a challenge for realism. Furthermore, the challenge thus created is fairly strong, as we cannot even attempt to address it in the ways holism has been traditionally addressed. For instance, we cannot try to confirm auxiliary hypotheses independently, by pairing them with different hypotheses to be tested (Ritson and Staley, 2021; Rowbottom, 2011)—because no mention of auxiliary hypotheses was made in the local version of holism.

However, there is a rejoinder which proceeds along somewhat similar lines and, as I said, comes from selective realism. This rejoinder says that those claims or posits are confirmed *along with* an encompassing hypothesis that survive theory-change, and can hence be confirmed beyond the degrees of prior belief through their involvement in *several distinct* overall hypotheses. As I shall argue, this effectively means embracing meta-empirical confirmation, thus enforcing the dilemma I mentioned earlier.

2 Structure of the Argument

2.1 From Local Holism and Underdetermination to a Challenge for Realism

To make things more precise, I assume the following local (hence the 'L') formulation of CH:

(LCH) The logico-semantic parts of a sufficiently general and substantive scientific hypothesis can only be confirmed via the confirmation of the hypothesis as a whole.

By 'logico-semantic parts', I essentially mean non-trivial entailments, i.e., meaningful claims the hypothesis inevitably commits us to. I shall make this clearer in the next section. The above is a *local* formulation as it only concerns one isolated hypothesis or theory, though on the right level of generality. However, note that, while the local formulation is *sufficient* to create a challenge for realism, it is by no means *necessary*: In order for the hypothesis in question to make contact with empirical evidence, it of course must be conjoined with assumptions about measurement-, background- and boundary-conditions (and so forth), and this only amplifies the problem.

In essence, I here offer three arguments that, taken together, present a serious challenge for realism. The first one is the following, where 'DEG' is short for 'degree':

(LCH) The logico-semantic parts of sufficiently general and substantive scientific hypotheses can only be confirmed via the confirmation of the given hypothesis as a whole.

(DEG) Provably, the degree of confirmation of the whole hypothesis is limited by the confirmation received by its parts.

∴ (C_1) If the confirmation of the whole hypothesis does not increase the degree of confirmation of one or more of its parts, the degree of confirmation of the whole hypothesis itself is bounded by the lowest prior probability over all those parts.

That is: LCH, combined with primitive facts about confirmation, has the implication that substantive scientific hypotheses might not be confirmable beyond the degree to which some of the theoretical posits involved in them are already believed. Note the interplay between the premises: If it wasn't for DEG, one might respond to LCH by saying: "Well, the hypothesis as a whole *can* be confirmed to any desired degree. There is no problem for realism."

Yet, C_1 is still merely a conditional claim that could be countered by a modus tollens: Hypotheses of the relevant sort receive lots of confirmation, so their parts *must* be confirmed along the way. Hence, further arguments are needed to move from conditional to unconditional claims. Furthermore, no connection was as yet made between (L)CH and UP.

Thus, consider the following, additional argument, where 'NTD' is short for 'necessary condition for holistic top-down confirmation':

(NTD) Increases in the degree of confirmation of an overall hypothesis can only increase the degree of confirmation received by any given part of that hypothesis if evidence does not speak at least equally in favor of a different hypothesis which does not contain that very part.

(UP) Typically, evidence does speak at least equally in favor of different hypotheses which do not contain all the same parts.

∴ (C_2) Typically, increases in the degree of confirmation of an overall hypothesis cannot increase the degree of confirmation received by some parts of that overall hypothesis.

NTD is a substantial assumption that will be discussed in more detail later, but on an intuitive level, it basically says that we cannot see a posit as confirmed along with an encompassing hypothesis if we can also do very well without the posit, by embracing an entirely different hypothesis. I believe this to be rather plausible. There now is an obvious challenge for realism, as C_1 and C_2 jointly imply that the degree of confirmation of general and substantive scientific hypotheses is indeed typically bounded by the lowest prior over all of their parts. On the level of the previous argument, the conclusion could be bypassed if confirmation was allowed to be transferred from whole to part, so that the dependency of the whole on its parts does not pose any serious restriction – both might simply be confirmed together. Here, underdetermination develops its bite, as it says that we will normally not be able to determine which parts (if any) *should* be seen as confirmed along with the whole.

If the above arguments are correct, we thus have reason to think that much of scientific confirmation ultimately hinges on matters of personal judgment, and we have no rational reason to believe what relevant hypotheses say about reality.[4] This I take to be the holistic challenge to realism, the one which depends on underdetermination, and which is established only by the two arguments in concert. It was here based on particularly weak assumptions: a local formulation of holism, standard facts from confirmation theory, and some very reasonable additional assumptions.

2.2 A Potential Rejoinder and a Dilemma for Realism

A reasonable rejoinder comes from *selective* forms of scientific realism (SSR). The central suggestion of SSR is to commit ourselves only to posits that are crucially connected to empirical success and sufficiently weak, or of the right kind, to have survived the switch from one theory to a rival successor. The relevant posits might be pieces of mathematical structure (Poincaré, 1907; Worrall, 1989), 'phenomenological' laws (Cartwright, 1983), or claims to the existence of some entity (Hacking, 1983). The underlying intuition always is the same: It is these very posits that are credible on account of the collective evidence for all the distinct, rival hypotheses in which they occur.

However, if LCH is true, then the posits in question cannot be *directly* confirmed in this way: The only way a given posit can be confirmed is by means of embedding it into a more encompassing hypothesis and transferring the confirmation of that hypothesis down to the part. But as argued above, this mode of holisitic top-down confirmation, as I have called it, is threatened by underdetermination. If this is correct, the only way forward for the SSRist is to invoke the very fact *that the posit in question has survived theory change* as evidence

[4]Emphatically, this must be sharply distinguished from such hypotheses' practical utility and overall empirical success in specific domains, which are excellent reasons for doing science and trusting in its instrumental value. Thus, like pretty much any 'anti-realist' philosopher these days, I suggest to remain enthusiastic about science, while still being skeptical about its ability to lead us to deep, mind-independent truths.

raising its credibility – and thus, a meta-empirical piece of information.[5] Hence, I shall also argue for the following, where 'MEV' stands for meta-empirical evidence:

(MEV) If the logico-semantic parts of sufficiently general and substantive scientific hypotheses can only be confirmed via the confirmation of the given hypothesis as a whole, SSR inevitably requires confirmation by meta-empirical evidence.

(LCH) The logico-semantic parts of sufficiently general and substantive scientific hypotheses can only be confirmed via the confirmation of the given hypothesis as a whole.

∴ (C_3) SSR inevitably requires confirmation by meta-empirical evidence.

C_3 is not a specifically anti-realist conclusion. In fact, defenders of meta-empirical confirmation could even argue that the success of SSR in explaining scientific success *en gros* supports their case. In turn, SRRists could bite the bullet and embrace meta-empirical confirmation. However, I believe this at least poses a dilemma for the SSRist: Abandon the approach, or side with an unlikely, maybe even unwanted, ally. My own suggestion here of course is that the controversial status of meta-empirical confirmation does in fact yield a case against SSR.

It is obvious that all central premises, LCH, NTD and MEV, need further support, and the bulk of the paper is devoted to developing this support. In fact, I shall recast NTD in a slightly different vocabulary and add some sensible qualifications. Furthermore, I will explain in more detail how the qualifiers of generality, substantiveness and scientificality support the ubiquity of the phenomenon described by LCH. In addition, it will be helpful to offer some actual examples, but space permits to discuss only a fairly limited sample. I shall hence appeal to three suggestive case studies that are alike in relevant respects and differ in others, so as to make a compelling case for the generality of the problem.

It will also be helpful to give some consideration to the meaning and justification of UP, as used in the second argument. I will here build on considerations by Stanford (2006) and Frost-Arnold (2019) in the context of their own arguments against realism; but as I will argue, the holism-based approach presented

[5] I will show later on why, under fairly modest assumptions, it is not possible to simply botch together the evidence received by the different, rival hypotheses in which the posit occurs without *ipso facto* embracing meta-empirical information as confirmatory.

here can employ the relevant cases of underdetermination in a way that Stanford (2006) and Frost-Arnold (2019) arguably cannot.

The paper is essentially organized into two parts: The first (Sect. 3–5) establishes how LCH and underdetermination create a challenge for realism, the second (Sect. 6–6.3) considers a response and establishes a dilemma for the realist. In detail, Section 3 justifies premise DEG, which follows straightforwardly from observations made by Schippers and Schurz (2020). Section 4 is concerned with the justification of premise LCH, which proceeds both in a principled way and by appeal to case studies. Section 5 then justifies premises NTD and UP and highlights the tension between them. Section 6 then discusses the rejoinder from SSR and responds by introducing the dilemma created by LCH. Section 6.3 finally vouches for the dilemma's anti-realist horn.

3 Confirmation of Hypotheses and the Confirmation of their Parts

3.1 Limits of Confirmation

Before getting to the real meat of the argument, it is necessary to justify premise DEG. For that, I need to appeal to a probabilistic, or broadly 'Bayesian', theory of confirmation. However, a serious caveat is in order here. It is customary to hold that a hypothesis h is confirmed by evidence e *just in case* $P(h|e) > P(h)$. But for reasons to become clear below, I will only accept one direction of this bi-conditional, namely, that cases of confirmation can be modeled as probability-increases; i.e., that when e confirms h, we get $P(h|e) > P(h)$.

Roughly, my reason for not seeing the probability increase as generally *sufficient* for confirmation relates to the fact that there could be other rationally compelling changes in doxastic attitude towards a given hypothesis, based on some piece of evidence or other, that do not satisfy all criteria for confirmation. Thus, since 'evidence' is at least partly ambiguous, I am here not assuming that all changes in credence based on evidence amount to confirmation, contrary to standard Bayesian confirmation theory.

That said, consider some primitive facts about Bayesian-style confirmation.[6] It is well known that pieces

[6] Throughout, I assume a first order language that allows variables $x, y, z, \ldots$, constant symbols $a, b, c, \ldots$, and function symbols $f, g, h, \ldots$ as terms of predicates, subject to the usual formation rules for formulae φ using the primitive logical symbols $\forall, \exists, \neg, \wedge, \vee$, and with $\varphi \rightarrow \psi$ defined by $\neg\varphi \vee \psi$.

of evidence that are mutually conditionally independent relative to h and its negation but confirm h (without entailing it) allow convergence to certainty in the long run.[7] Thus, we may reasonably become more and more confident about a given hypothesis by gathering more and more independent pieces of evidence.

However, it also holds that any implication of some hypothesis h,

> call it c, that is not confirmed by any one of the evidences e_i is sufficient to prevent convergence to certainty. Since c's probability is not raised by any of the e_i it holds that $P(c|e_1 \wedge e_2 \wedge \ldots) = P(c)$. But $P(c|e_1 \wedge e_2 \ldots)$ is an upper bound of $P(h|e_1 \wedge e_2 \ldots)$ (since h entails c), whence $P(h|e_1 \wedge e_2 \wedge \ldots)$ is forced to stay below the low value of $P(c)$ even if $n \to \infty$. (Schippers and Schurz, 2020, 332–3; notation adapted)

This is possible since we do not automatically get $P(c|e) > P(c)$ in case $P(h|e) > P(h)$ (see Appendix, A 1). Now, there is a rather trivial variant of this issue which could be seen as a general criticism of Bayesian confirmation theory: For any given h and any given piece of evidence e_i, $h \vee \neg e_i$ is an implication of h that is provably not confirmed by e_i (see Appendix, A 2).[8] Hence, if all logical implications are considered, any piece of evidence is trivially ruled out as confirming h.

3.2 Towards a Non-Trivial Challenge

To make the issue non-trivial, I will embrace a notion of *relevant* entailment here, as has proven useful in several distinct debates (Roche and Sober, 2017; Schippers and Schurz, 2020; Sprenger, 2011). Following the particular approach by Schippers and Schurz (2020), thus call some statement φ a 'content element' of some hypothesis h if it is one of the most primitive relevant entailments of h, and call a conjunction of content elements a 'content part'.[9]

[7] That is: If there are e_i $(1 \leq i \leq n)$ such that $P(e_i|(\neg)h \wedge e_j) = P(e_i|(\neg)h)$ for $i \neq j$ and $P(h|e_i) > P(h)$ for each i, then $\lim_{n\to\infty} P(h|e_1 \wedge \ldots \wedge e_n) = 1$.

[8] I was made aware of this fact by G. Schurz (private communication). More general theorems and proofs are found in Popper and Miller (1983), although they are ultimately concerned with a different issue.

[9] More precisely, φ is a content element of h exactly if (a) $h \models \varphi$, (b) φ is elementary in the sense of not being logically equivalent to a conjunction $\varphi_1 \wedge \varphi_2 \wedge \ldots \wedge \varphi_n$ $(n \geq 1)$ where each φ_i has a smaller number of primitive symbols when φ is expressed in 'negation-normal form' – only $\neg, \wedge, \vee$ as connectives, and $\neg$ in front of propositional formulae by *de Morgan* (Schurz and Weingartner, 2010, 425) – and (c) φ is relevant to h in the sense that no predicate in φ is replacable on some of its instances by another predicate of the same arity *savla validitate* of $h \models \varphi$ (Schurz, 2014, 93). Thus building on entailment means following an approach to content that goes back to Carnap (1937), which was intended as an account of logical content only. However, the issue at best remains unaltered and at worst becomes amplified when we also add in considerations of non-logical content. Furthermore, Schippers and Schurz (2020) call 'evidence-transcending' those parts that are not implied by the evidence. Since all pieces of evidence considered below are evidence-transcending in this sense, I will largely omit the qualifier.

Building on this notion of content, Schippers and Schurz (2020) hold a hypothesis h to be *genuinely* confirmed by evidence e only to the extent that e increases the probability of some of h's content parts φ, while not also decreasing the probability of others.[10] And in case e confirms all content parts φ, it may be said to *fully* genuinely confirm h.

The trivial sense of unconfirmability mentioned above is thus avoided.[11] However, as Schippers and Schurz (2020, 332) are well aware, this does not resolve the issue created by the dependence of a hypothesis' confirmation on the confirmation of its *non-trivial* implications: "a necessary condition for convergence to certainty is [full genuine confirmation]. Only one [...] content part of h [...] that is not confirmed by any one of the evidences e_i is sufficient to prevent convergence to certainty." Much of what follows builds on this insight.

4 Local Confirmational Holism

4.1 Substantive Hypotheses

What is the core message of CH? Quine (1951, 38) famously put the message thus: "statements about the external world face the tribunal of sense experience not individually but only as a corporate body". Fodor and Lepore (1992, 41) put it more succinctly as follows: "every statement in a theory (partially) determines the level of confirmation of every other statement in the theory". But why should this be so?

Recall that I formulated my version of holism, LCH, with respect to a single hypothesis of a certain kind. Specifically, a hypothesis vulnerable to the present line of argument was supposed to be 'general', in the sense that it cannot be intended as a hypothesis about a particular set of phenomena, but must rather concern a broader range of distinct phenomena. Hypotheses of the former sort might be immune to the problem I'm embarking upon, but they may as well be thought of as 'merely descriptive' or 'fully empirical'. They are hence also not the kinds of claims of interest to scientific realists.

Furthermore, 'substantive' means that the hypothesis in question cannot be considered just a useful tool or

[10] That is, e at least partially genuinely confirms h exactly if $P(\varphi|e) \geq P(\varphi)$ for all content parts φ of h and $P(\varphi'|e) > P(\varphi')$ for at least some of h's content parts, φ'.

[11] I.e.: We can replace predicates in the e_i-part of $h \vee \neg e_i$ *salva validitate* of $h \models h \vee \neg e_i$. See footnote 16.

formalism whose content and interpretation do not matter: It must be taken to be 'about' reality. The qualifier 'scientific', of course, implies that the hypothesis in question must be formulated in a sufficiently precise language so as to even allow empirical testing at all; that is, it must have discernible, typically quantitative, and at least probabilistic, consequences, thus putting itself at the risk of being wrong (Vickers, 2019; Douglas, 2009).

In the following, I will discuss three distinct examples, to suggest that the phenomenon described in LCH is widespread and concerns much of science. I will do this by building on essential differences between features that might be suspected to be the real culprit instead, such as belonging to a specific discipline or being of mathematical form; though I admit that the case is most compelling or most easily seen in highly mathematicized disciplines. I will finally also go into the principled reasons for why I believe LCH to be a widespread phenomenon.

4.2 Example: The Maxwell Equations

Hypotheses of the relevant guise usually figure at a relatively fundamental level of theorizing. A nice example are Maxwell's equations in Classical (CED), and even Quantum Electrodynamics (QED). At bottom, these equations put forward the hypothesis that there are (at least two) fields out there that relate to each other in a particular way. The fields in question are the electromagnetic field-strength tensor, $f^{\mu\nu}$, and the current-density, j^{ν}, which are discernible from one another by means of their distinct properties, expressed in terms of different mathematics. The relation between them is expressed by the differential equations.

Roughly put, the hypothesis in question is thus of the general form

$$\exists f \exists j (F_1[f] \wedge F_2[j] \wedge R[f, j]),$$

where F_1 and F_2 specify the fields' properties, and R specifies the relation(s) expressed by the Maxwell equations.[12] But then $\exists f \exists j (F_1[f] \wedge F_2[j])$ is a content-part of this hypothesis, meaning that acceptance of the Maxwell equations commits us to the assumption that there are such fields as specified by F_1 and F_2. How could we possibly confirm this, other than by drawing on the hypothesis as a whole?

[12] $X[\cdot]$ denotes a formula in which the bracketed variables occur freely.

On the pains of introducing further notation, it is helpful to apply a little more rigor here. In standard presentation, Maxwell's equations are usually given as

$$\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x),$$
$$\varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0,$$

where $f_{\gamma\delta} = g_{\nu\delta} f^{\mu\nu} g_{\gamma\mu}$, $\varepsilon^{\mu\nu\gamma\delta}$ is the Levi-Civita symbol, and $\partial_\mu f^{\mu\nu}(x)$ implies summation over μ. However, without any quantifiers, these systems of equations are actually open formulae, not claims at all.

A little more carefully, the hypothesis in question should thus be stated as follows. Let $M = \langle \mathcal{M}, g \rangle$ be Minkowski spacetime, and let $\Gamma^1_{\langle n,m \rangle} := \Gamma^1(T^n{}_m\mathcal{M})$ the set of once-differentiable tensor-fields of rank $\langle n, m \rangle$.[13] Then the hypothesis of interest is:

$$\exists f \in \Gamma^1_{\langle 2,0 \rangle} \exists j \in \Gamma^1_{\langle 1,0 \rangle} \forall x \in \mathcal{M} \forall 0 \leq \mu, \nu, \gamma, \delta \leq 3 (\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \wedge \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0) \tag{Maxwell 1}$$

We can unpack the quantification restricted to the given sets as follows, which makes it explicit that existence claims regarding certain fields form part of the hypothesis:

$$\begin{aligned}&\exists f \exists j \forall x \forall \mu, \nu, \gamma, \delta (f \in \Gamma^1_{\langle 2,0 \rangle} \wedge j \in \Gamma^1_{\langle 1,0 \rangle} \wedge \\ &\qquad (x \in \mathcal{M} \wedge \mu, \nu, \gamma, \delta \in \{0, 1, 2, 3\} \rightarrow (\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \wedge \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0)).\end{aligned} \tag{Maxwell 2}$$

Furthermore, if we pull the universal quantifiers past the first two conjuncts, we here recognize the same logical form as in the much coarser reconstruction given above, where the first two conjuncts within the scope of the existential quantifiers correspond to the non-relational formulae ($F_i[g], g \in \{f, j\}$) and the

[13] That is, $\mathcal{M} \cong \mathbb{R}^4$, $g = \mathrm{diag}(-1, 1, 1, 1)$ the (Lorentzian) metric, and

$$\mathcal{M} \ni x \mapsto T^{\mu_1 \ldots \mu_n}{}_{\nu_1 \ldots \nu_m} \in T^n{}_m\mathcal{M} = \cup_{x \in \mathcal{M}} (T_x\mathcal{M})^{\otimes n} \otimes (T^*_x\mathcal{M})^{\otimes m},$$

with $T_x\mathcal{M}$ the tangent space of $\mathcal{M}$ at x, $T^*_x\mathcal{M}$ the dual of $T_x\mathcal{M}$, and the space-time derivatives $\partial/\partial x^\rho$ of $T^{\mu_1 \ldots \mu_n}{}_{\nu_1 \ldots \nu_m}(x)$ well-defined at any x.

remainder to the relational formula ($R[f, j]$). Obviously,

$$\exists f (f \in \Gamma^1_{\langle 2,0\rangle}) \wedge \exists j (j \in \Gamma^1_{\langle 1,0\rangle}), \qquad (c)$$

the claim that there are things out there which are well-described as certain tensor-fields on a Minkowskian spacetime, is a content-part of (Maxwell 2).

For any empirical evidence, c is certainly evidence-transcending in the 'right' sense of not being entailed by the evidence (Schippers and Schurz, 2020, 330). However, in order for Maxwell's equations to count as substantive, it must be presupposed that (Maxwell 2), and with it c, is neither to be interpreted purely in mathematical terms, nor as an 'empty calculus' which can be used and interpreted in various ways. Rather, it must be assumed that the chosen mathematical representation, c, refers somewhat adequately to reality. But when taken as substantive in this sense and considered *in isolation*, how could a claim such as c be brought into contact with evidence at all? In other words: is c, by itself, not also be 'evidence-transcending' in the problematic sense of being *beyond contact* with any possible evidence?

To see the point clearly, consider a concrete, simple way of incrementally confirming Maxwell's equations. Thus, imagine some measurement near a wire in which a current flows, with outcome m, specifying some sensible experimental average over several runs. Assume now that (a value close to) m is predicted by the Biot-Savart law,

$$\boldsymbol{B}(\boldsymbol{x}) \propto \int \frac{\boldsymbol{j}(\boldsymbol{x}') \times (\boldsymbol{x} - \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|^3} \, \mathrm{d}^3 x' \, ,$$

where $\boldsymbol{B}$ is the magnetic field vector specifying the magnetic components of $f^{\mu\nu}$ in the laboratory frame, and $\boldsymbol{j}$ contains the three-vector components 1–3 of j^ν in that frame, which provide the measured current.

Since this equation follows from (Maxwell 2) in the special case where $f^{\mu\nu}$'s electric components are static in the given frame, we may hold this success of the Biot-Savart law to confirm (Maxwell 2), including c. But the crucial point is that c *in isolation* has no such implication.

Of course, a relevant sort of measurement, m, need not be *implied* by the hypothesis under scrutiny in order to have confirmational impact. It merely has to reasonably increase the hypothesis' probability. So still considering c in isolation for the moment, there are three basic possibilities how this could happen: (i) c might, indeed, imply m, and then it is a theorem that $P(c|m) > P(c)$; (ii) m might raise the probability of c directly; or, (iii), the likelihood $P(m|c)$ of observing m given c might be greater than $P(m)$, so by Bayes' theorem, $P(c|m) > P(c)$.

However, remember that we are now asking for confirmation of c *alone*, not *via* the confirmation of (Maxwell 2). Under these conditions, (i) can be dismissed right out of hand: The mere assumption that there are certain kinds of fields out there – *full stop* – implies nothing about any measurement. To see this, imagine someone telling you in an uninitiated conversation that there are certain fields out there that satisfy the relevant mathematical representation, and that this implies that a specific current will be measured in some wire. Immediately, you will assume that the person has a lot of theoretical background knowledge and is aware of deductive relations from c *conjoined with that background knowledge* to the measurement claim – unless you assume they are simply crazy. Furthermore, nothing hangs on the particular example: by analogous reasons, no implication from c *alone* and *any other* kind of measurement can be established.

However, closer scrutiny reveals that pretty much the same is true about (ii): Assume that the person just tells you that a given measurement makes it probable that the fields exist. Immediately, you would be prone to wonder based on what background knowledge the person has estimated the probability. If we disallow this appeal to further background knowledge, such as the relation between c and the Maxwell equations—i.e., absent further theoretical assumptions about what $f^{\mu\nu}$ and j^{ν} *do*; and specifically, how they interact with relevant measuring devices—it seems entirely unreasonable to just stipulate some direct probabilistic connection.

So the only *prima facie* plausible option is (iii), which constitutes the standard reasoning chain employed in Bayesian inference methods (Brimacombe, 2018, §1.2; Lawrence, 2019, §7.2). Likelihoods are usually inferred either directly from the hypothesis, h, in question, or defined by folding in a suitable, probabilistic model of the measurement under the assumption that h is true (Gillies, 2000). For instance, if h specifies an elementary probability p for a certain outcome, we model the measurement as being binary, and the variables measured as being independent and identically distributed, then the likelihood for observing $m = k/n$ should be chosen to be of binomial form. Similarly, if h is deterministic, m specifies an expectation value, and the number of observations is very high, one may use the central limit theorem to approximate the likelihood as a normal distribution, centered on the value predicted by h, and with standard deviation reflecting the experimental error as given by the model of the measurement (see Bolstad and Curran, 2017).

But again we must ask: What choice of likelihood could be justified on the basis of the mere assumption that there are certain fields out there in spacetime, well-represented by tensors on a Minkowskian manifold? Remembering that we are still considering c in isolation here, I see only three responses: Either, (a), we conclude that it is impossible to define any likelihood at all, and hence our Bayesian updating cannot get off the ground. But *ispo facto*, this means that we cannot even begin to confirm c in isolation.

Alternatively, (b), we may try to evaluate the probabilistic implications of c regarding m. However, since c does not constitute a probabilistic hypothesis, it seems that the inference to a likelihood directly from c is equally bound to fail, for almost the same reasons as this was the case with (ii) above: Why should the mere assumption that there are certain fields out there impact our expectations regarding certain lab-operations leading to some outcome m?

Finally, (c), we may consider how a probabilistic model of some measurement, sensibly related to c, may lead to the definition of a likelihood for outcomes m, in the ways discussed above. However, this approach is easily seen to suffer from the same basic deficiencies as any more direct route: Without any clues as to what $f^{\mu\nu}$ and j^{ν} do, how could we use the bare postulate of their existence to make up our minds about the statistics of a purported measurement with result m? It seems that there is just no justifiable way to connect m probabilistically to c, without further theory about the behavior of the fields thereby postulated.

Form the foregoing, we may hence conclude that, without the 'bare' Maxwell equations (plus boundary conditions, a model of the measurement, and so forth), our credences about m should not be altered by the assumption that c. Hence, under these conditions, Bayes' theorem yields

$$P(c|m) = \frac{P(m|c)}{P(m)}P(c) = P(c).$$

Note, crucially, that the same observations apply to the bare equations without c, or a suitable replacement thereof: The mere assertion that there is 'something' which satisfies certain relations ($\exists f, jR[f, j]$), without saying more about what that something is, is insufficient to massage the relating equations into numerical predictions, determine expected experimental frequencies, or form other (probabilistic) expectations that sensibly relate to the evidence. This is holism alright: "statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (Quine, 1951, 38), and "every statement in a theory (partially) determines the level of confirmation of every other statement in the theory" (Fodor and Lepore, 1992, 41).

4.3 Beyond Physics

A first distraction should be dispelled here: Duhem (1914, 180 ff.) famously thought that these observations only pertain to physics, and the reader may be misled, by my choice of example, into thinking that I concur. Hence, consider another example from a different discipline, namely the Price equation from evolutionary biology,

$$\Delta \bar{z} = \mathrm{Cov}\left(\frac{w_i}{\bar{w}}, z_i\right) + \frac{\mathrm{Exp}(w_i \delta_i)}{\bar{w}}.$$

Here, i is "a label for any sort of property of things in the [...] ancestral population", w_i is "the fitness of the ith type", $\bar{w} = \sum_i q_i w_i$, where q_i is the frequency of i in the ancestorial population, and z_i "can be a measurement of any property of an entity with label, i", such as "the frequency of a gene, the squared deviation of some phenotypic value in relation to the mean, the value obtained by multiplying measurements of two different phenotypes of the same entity and so on" (Frank, 2012, 1005). Finally, $\delta_i = z_i' - z_i$, where z_i' is "the average measurement of the property associated with z among the descendants derived from ancestors with index i", and $\Delta \bar{z} = \bar{z}' - \bar{z}$ where $\bar{z}$ is defined analogous to $\bar{w}$, and $\bar{z}' = \sum_i q_i' z_i'$, with q_i' "the fraction of the descendant population that is derived from members of the ancestral population that have the label i" (*ibid.*).

This equation has been recognized for formalizing fundamental Darwinian intuitions (Gardner, 2020) and for playing a similarly fundamental role in evolutionary biology as do Newton's laws in classical physics

(Luque and Baravalle, 2021). Hence, this is a hypothesis at the right level of generality. However, as such, the Price equation does not make any predictions; for that it has to be conjoined with causal models of the dependencies in question (Okasha and Otsuka, 2020), not entirely unlike the force models one has to supply to the Newtonian equations in order to retrieve something testable (Giere, 2010).

Using such a causal model, it is possible to retrieve a variant of the Price equation called the breeder's equation, which is an approximation that builds on independence assumptions that can only hold asymptotically. As with the frame-relative static field-assumption and the Biot-Savart law in relation to the Maxwell equations,

> to the extent that those assumptions approximate reality, it [the breeder's equation] allows us to make inferences about unobserved or unobservable evolutionary changes under future or hypothetical conditions. (Okasha and Otsuka, 2020, 7)

Thus, with an appropriate model of the observational conditions, such facts as that this equation "has long served for breeders to predict a change in the mean phenotype before performing artificial selection" (Okasha and Otsuka, 2020, 7) may be held to confirm the breeder's equation and, by extension, the general Price equation.

To treat the Price equation on the same footing as Maxwell's equations, we would have to explicitly introduce quantifiers ranging over well-defined sets. For reasons of brevity, I'll eschew a thorough reconstruction at this point though, and jump right to the conclusion: Using more rigor, we would have to commit ourselves to a particular mathematical representation of, in particular, the fitness variable w_i, subject only to $w_i = \bar{w}(q_i'/q_i)$. This commitment can be cashed out as an existence claim that, in turn, can be tested along with, but not in isolation from, the entire equation. So an analogous game as above may be played in this case as well, even though the quantities in question are certainly less 'esoteric' than tensor fields.

4.4 Beyond Mathematics?

Both examples I have investigated were thoroughly mathematically formalized. So, was Duhem (1914, 180) at least right that hypotheses wherein "mathematical theory has not yet introduced its symbolic representations" are in general immune to the challenge? I am not convinced that this is correct, though I admit that the issue can be most straightforwardly seen in the case of fully formalized examples.

To see why it might not only be mathematical hypotheses that matter, consider plate tectonics as an example, i.e., the hypothesis that the earth's crust is made up of large, rigid, rocky plates that move on top of a layer of molten rock, causing such things as continental drift and earthquakes. This hypothesis has been considered as a prime example of a scientific hypothesis that is hard to disbelieve (Fahrbach, 2021, 286; Magnus, 2013, 50), in contrast, say, to theories in fundamental physics. And *prima facie*, it could be clearly stated without the use of any mathematics.

Can we raise a similar challenge for plate tectonics as for the Maxwell and Price equations? The answer is yes: If plate tectonics is presented as above, it has as a content-element the claim that the earth's crust is made up of large, rigid, rocky plates, and it is clearly impossible to gather evidence about these plates without making reference to their behavior, such as their movement relative to one another. Thus, the jigsaw puzzle-like look of the continents, or the distribution of earthquakes and volcano eruptions across the globe, all count as evidence for the existence of tectonic plates only *via* the full hypothesis that also concerns their motion. I believe this provides reason to think that LCH is rather wide-spread.

However, even assuming that one is less convinced by this example than by the other two – as the sharpness of the challenge obviously withers with the sharpness of the example – this does not make the challenge any less daunting, for two reasons.

First, there is reason to think that hypotheses that do receive respectable support from evidence not only make manifold predictions, but also predictions precise enough so as to be 'risky', in the sense of coming out wrong under a large range of scenarios (Vickers, 2019). Making (many) such risky predictions requires sufficiently precise (and abstract) concepts, which are typically the subject of mathematics. For instance, Vickers (2013, 196) discusses such things as the correspondence between prediction and experiment "to many significant figures" across several distinct types of experiment as potentially defining a "degree of impressiveness" of a (novel) prediction.

Thus, mathematics might need to enter at some point for serious confirmation to be possible: Just think of the discussion on adaptationist explanations as mere 'just-so stories' (Gould and Lewontin, 1979) and Orzack and Sober's (1994) response in terms of – fully mathematicized – optimality models' predictive success. Or think of the ways in which social scientists formulate hypotheses about dedicated test statistics in order apply mathematical statistics to test them. Of course, the same is true of plate tectonics: "To be a [...] testable [...] model, plate tectonics requires spherical geometry and spherical kinematics in terms of finite rotations conveniently parametrized by their angle and axis and described by unit quaternions." (Schaeben et al., 2024, 4469)

Second, when it comes to the details of confirmation, another concern emerges: Oreskes (2003, 21) and Manson (2003, 31) describe the early stages of the gradual acceptance of plate tectonics in the 1950as and 60s in virtue of then-novel evidence in its favor. Part of the most surprising evidence at the time were reversed magentizations in certain rocks, due to a differently oriented exposure to the gravitational field, as well as corresponding magnetic 'stripes' on the sea floor. However, interpreting these pieces of evidence requires the use of CED – and with it, the Maxwell equations.

This is clearly not the same challenge as before (LCH), but rather an appeal to traditional views on holism – that "[t]he unit of empirical significance is the whole of science" (Quine, 1951, 42), or at least that we "can never subject an isolated hypothesis to experimental test, but only a whole group of hypotheses" (Duhem, 1914, 187). However, I explicitly never opposed broader appraisals of holism in arguing that LCH is *sufficient* to create a problem. Furthermore, as we can see, the hypothesis ultimately relied upon is the prime example I had used to explain the occurrence of LCH, and of course electromagnetic theory is required for evidence-generation in a fair bit of science. Hence, even if one rejects LCH *per se* for sciences that are not (yet) mathematically formulated – which I believe one cannot –, it has far-reaching consequences.

4.5 Principled Reasons for LCH

The presence of LCH across the sciences, as I have tried to convey it above, is not so much due to disciplinary peculiarities or the use of mathematics, but can be traced to the qualifiers of generality and scientificality I had introduced earlier.[14] For, generality implies that one entertains concepts that are abstracted away from

[14]Of course, without the substantiveness-qualifier, there couldn't be any challenge for realism.

individual observations to such a degree that they can be related to other, often quite different observations, or are even generated fully independently of any concrete observation whatsoever. Scientificality implies that the concepts in question are precise enough to allow precise and testable predictions, when conjoined with further elements such as relations (equations) they enter into. At the point when both these conditions are fulfilled, one has moved away from what is observable by empirical means so far that any *immediate* connection between individual commitments made and possible experimental outcomes breaks down. This was the upshot of the previous sections.

For example, it is one thing to claim that there is an area in the proximity of every one out of a certain range of similar solid objects (magnets) in which the other objects from that range react by being attracted or repulsed, depending on orientation: This hypothesis can very reasonably be said to make it more likely that the next object out of the given range will react in the same way. It is a whole other thing to claim that there are extended, continuous objects distributed across a space-time continuum that are well-described by the association of mathematical tensors to points in that continuum.

But clearly, we cannot move from the former sort of hypothesis (though, in the right theoretical context, certainly from the latter) towards an account of what goes on when connecting a phone to a power outlet. Furthermore, could we extend such a hypothesis in such ways that it encompasses also these latter phenomena without running the risk of becoming vulnerable to the challenge raised above? This seems doubtful. Insofar as all this is correct, LCH should be considered a fairly general claim that pertains to much of science.

5 Underdetermination

5.1 Holistic Top-Down Confirmation

There is a reasonable response to the foregoing, which I coined 'holsitic top-down confirmation' above: One might counter that scientific practice shows that some content-parts c are typically confirmed along with the encompassing hypothesis h.

This response is congenial to the very move at the heart of all forms of realism (e.g. Boyd, 1973, 2–3): Because some theoretical posit has figured successfully in the (probabilistic or deductive) prediction of evidence, we are justified in committing ourselves to said posit. However, we already know that promoting $f^{\mu\nu}$ to an operator-valued function as in QED, say, saves the predictions as well. Similarly, the assumption of perfectly rigid tectonic plates can, and should, be replaced by one which allows such things as creep in the plates (Keith, 2001, 237–8). But this says that we must be careful in transferring confirmation from the top down: Just because a posit has been useful in generating predictions, this does not merit realist commitment to it (also Vickers, 2013).

We might try to follow a line of argument by Okasha (2002) to avoid the daunting consequences at this point. Thus, Okasha (2002, 309–10) argues that CH can sometimes be invoked to *defeat* UP, as the relation of one theory to a more encompassing one into which it is embedded can make a difference. For example, before the special theory of relativity was recovered as a limiting case of the general theory, it could be seen as empirically equivalent to the Lorentz-Fitzgerald ether theory. But after that recovery, there was no question about which of the two was preferred anymore.

We might very similarly respond that (L)CH actually speaks *against*, say, the underdetermination between c and some suitable replacement – call it q – that involves quantum operators. But this doesn't quite achieve what we want it to: The contrastive argument may be correct; we may, indeed, get a relative confirmation of a corresponding quantum-version over (Maxwell 2), simply because (Maxwell 2) doesn't receive any confirmation from certain evidences e at all. We may even have $P(q) > P(c)$, taking into account that priors are always relative to a knowledge base and that this base may also include the successes of non-relativistic quantum mechanics. But note that none of this directly answers the question whether the evidence really confirms q. This remains as much of a question as this was the case with c and the evidence for classical electrodynamics (see also Salmon, 1990, 196, in this connection).

What is needed to safeguard against holism's impacting realism is a concrete justification for holistic top-down confirmation. So what could possibly justify the transfer of confirmation from hypothesis to content-part? Two necessary conditions for this have been considered by Schurz (2022, 8; notation adapted), according to which we can only have $P(d|e) > P(d)$ for some content-element d of an h that satisfies $P(h|e) > P(h)$, if:

(1.) d is necessary within h to make e probable, *i.e.*, there exists no conjunction h^* of content elements of h that makes e at least equally probable ($P(e|h^*) \geq P(e|h)$) but does not entail d, and

(2.) it is not the case that d is an existential quantification, $d \equiv \exists x h(x)$, and h results from fitting the value of x in $h(x)$ towards e, such that an equally good fitting of $h(x)$ would have been possible for every possible alternative evidence e'.

Thus, neither can d be predictively superfluous to the evidence, nor can it be a foil for *ad hoc* hypotheses designed to accommodate arbitrary evidence. However, these are only necessary conditions: Their failure should immediately forestall holistic top-down confirmation, while the fact that both obtain does not automatically justify it. Furthermore, it is unclear that they are *jointly* sufficient, and that there are no further constraints that should be imposed on holistic top-down confirmation.

Here is, hence, a suggestion for an additional necessary condition, which is a more precise formulation of the consequent of premise NTD from Section 1:

(3.) There is also no alternative hypothesis $h^\dagger$ such that $P(h^\dagger|e) > P(h^\dagger)$, for any e that satisfies the above conditions, and d is not a(n evidence-transcending) content part of $h^\dagger$.

In other words, I suggest that we should also require that there be no alternative hypothesis that does not draw on d but is confirmed by (at least) the same evidence. This covers cases where some posit contained in an old hypothesis turned out to be superfluous for getting the evidence right; the kinds of cases anti-realists have traditionally appealed to and the kinds we have discussed above (non-rigid plate tectonics, QED and so on).

But this is as yet a little too strong, for maybe the evidence *does* confirm $h^\dagger$, though not as strongly as h.[15] Under these conditions, we might hold d to be confirmed along with h after all. Hence, a more careful version of NTD's consequent should read as follows:

(3.') There is also no alternative hypothesis $h^\dagger$ such that $P(h^\dagger|e) > P(h^\dagger)$, where the confirmation received by $h^\dagger$ from e is at least as great as that received by h, for any e that satisfies the above conditions, and d is not a(n evidence-transcending) content part of $h^\dagger$.

[15] I have been confronted with this objection by an anonymous referee.

The condition of greater or equal confirmation could be formalized in several ways; e.g., by means of the difference measure $P(h^{\dagger}|e) - P(h^{\dagger}) \geq P(h|e) - P(h)$. The dependence on priors would then introduce a subtlety, but I believe it is fairly reasonable to assume that, in the relevant historical cases, the then-rejected hypothesis always enjoyed at least the same initial plausibility (relative to the same set of evidence) as the novel one.

Indeed, I suspect that pretty much all historically interesting cases of interest are cases which violate (3.') as well: In the case of CED and QED, all known pieces of evidence in support of CED should be as much supportive of QED. Furthermore, QED predicts many pieces of evidence that CED doesn't, and so is confirmed by these to a greater degree. Accordingly, most scientists and philosophers would concur that electromagnetic phenomena turned out not to be, in fact, produced by 'classical fields'. Similarly, we would probably say in retrospect that experimentation by Priestley confirmed combustion by means of oxygen just as much as (or even more than) the existence of phlogiston, whereas additional evidence only supported the oxygen hypothesis. And finally, soft-plate tectonics can accommodate additional evidence over traditional rigid-plate tectonics, even though rigid-plate tectonics is still referred to as the 'standard theory' in virtue of its successes (see Ribeiro and Mateus, 2012). In all cases, things erroneously appeared to speak in favor of the respective now-rejected posit at first.

5.2 Whence Underdetermination?

The foregoing hints at the relevance of (transient) underdetermination: It usually takes some time until the available evidence very clearly favors a novel hypothesis over an old one. Hence, the same evidence may support several rival hypotheses which do not have all the same contents, at least for certain periods of time. However, what are reasons to think that this is 'typical', as claimed by premise 'UP' in the second argument considered in Section 1? I will here lean on work on the historically motivated versions of anti-realism by Stanford (2001, 2006) and Frost-Arnold (2019); though, emphatically, I am not in the business of discussing or reconstructing either Frost-Arnold's or Stanford's arguments in detail. Nor am I claiming that either author invokes holism. Rather, I will cherry-pick on relevant parts of their respective overall argumentations, and the reasons for underdetermination they each rely on, to support my own case.

As is well known, Stanford (2001, 2006) in essence puts forward an inductive argument to the effect that we should expect relevant rivals to our present theories to be discovered in the future. This Stanford calls 'the problem of unconceived alternatives', and characterizes the historical pattern supporting it as a pattern of "recurrent, transient underdetermination" (Stanford, 2006, 17).

Notably, Stanford uses scien*tists*' failure to conceive of relevant alternatives in the past as an induction base, not scientific theories' failure to be (even approximately) true (Laudan, 1981). Thus, his argument is not vulnerable to the same criticisms as is the classic pessimistic induction (see Magnus, 2010, 814). However, in order to turn this into an actual argument against realism, Stanford (2006, 29) appeals to *eliminative* reasoning:

> the historical record suggests that in science we are typically unable to exhaust the space of likely, plausible, or reasonable candidate theoretical explanations for a given set of phenomena before proceeding to eliminate all but a single contender, but this is just what would be required for [...] eliminative inferences to be reliable.

Hence, the eliminative inferences scientists appear to make, and realists seem to endorse, are unsound, and so we have no good claim to present theories being approximately true on account of these inferences. However, it is unclear in how far this pattern of eliminative reasoning is even involved in science (Magnus, 2010, 816), or in how far a defense of realism requires it to be (Psillos, 2009). Furthermore, Frost-Arnold (2019, 914) has recently pointed out that Stanford cannot appeal to some of the "'greatest hits' from the antirealist catalog", as these do not fit the pattern identified by Stanford.

Accordingly, Frost-Arnold (2019, 910) identifies a different problem that he calls 'the problem of misleading evidence': "The total body of evidence used by scientists at a particular time was often unrepresentative or otherwise misleading." For example, consider the case of the spontaneous generation of life from inanimate matter, as discussed by Frost-Arnold (2019, 912): Observing boiled and carefully sealed gravy under a relatively high-quality microscope, Needham and Buffon could see bacteria in motion, in contrast to Spallanzani and others, who used less powerful instruments. However, Needham and Buffon concluded from this that these bacteria were alive and, due to the proper concealment, their life had to have been created spontaneously. From the present vantage point, however, this phenomenon is instead explained by Brownian motion (Frost-Arnold, 2019, 912).

Frost-Arnold does not spell out explicitly what he means by 'otherwise misleading';[16] but I suggest that one plausible reading is that evidence is *ambiguous*. Or in other words: It isn't always perfectly clear *what the evidence even is*. If this is the case, certain pieces of evidence may appear to support a given hypothesis, where in reality they don't.

For example, the episode recounted above might be reconstructed as follows: Assume that the hypothesis of spontaneous generation of life from inanimate matter, s, entails (or is supported by the fact) that bacteria in motion can be observed in boiled and concealed gravy, β. Then we have $P(s|\beta) > P(s)$. However, s also has the more specific implication that these bacteria cannot always be moved by external (or at least: non-vital) forces, $\neg\varepsilon$, including external molecular forces consistent with Brownian motion: At least sometimes, it has to be the spontaneously generated life that accounts for the motion.

Now, from the present vantage point, what Needham and Buffon observed was not really β but β': a *specific* pattern of motion, consistent with having its origins in external molecular forces as described by Einstein (1905). We may thus assume that $P(\beta'|\varepsilon) \gg 0.5$, i.e., that the external force-hypothesis makes *this kind of* pattern rather likely. Further, if $P(\varepsilon) \leq P(\beta')$ or at least $P(\varepsilon) \approx P(\beta')$, which seems plausible given that the external force hypothesis was disbelieved at the time, we get $P(\varepsilon|\beta') \gg 0.5$ by Bayes' theorem. However, since the confirmation of s is bounded by the confirmation received by its consequences, $P(s|\beta') \leq P(\neg\varepsilon|\beta') = (1 - P(\varepsilon|\beta')) \ll 0.5$. So under these assumptions, s cannot receive any confirmation from the evidence *as conceived of today*. Notably, the relevant evidence was not of low but instead of rather high quality, and it still misled Needham and Buffon into believing they had confirmed spontaneous generation.

How do these considerations support UP, as used in the second argument of Section 1? Both Stanford (2006) and Frost-Arnold (2019) identify historical patterns that support the view that underdetermination occurs regularly, in the sense required by premise UP. But crucially, both offer *distinct reasons* for this: On the one hand, humans are arguably bad at anticipating future theoretical developments; so the *posits* needed to

[16] Frost-Arnold (2019, 910) explicitly refrains from offering a general definition. He considers a potentially useful one – that some evidence incrementally confirms a theory which is not even approximately true – but also acknowledges its limitations, as pointed out by Fallis and Lewis (2014). As should be clear from what follows, I also reject this as a general definition of evidence's being misleading.

accommodate further evidence are never dictated by evidence itself. On the other hand, *evidence* is generally in need of interpretation, and interpretation is rarely straightforward – a rather Kuhnian point. So what we take the evidence to be, and what we take it to support, may change over time. Evidence alone is thus powerless to determine which hypotheses we should believe, in at least two distinct ways: Both the posits needed to accommodate all available evidence, and the best interpretation of that evidence, are underdetermined by the bare evidence itself. Jointly, this makes it rather plausible that underdetermination should occur rather frequently (now as then), and that its appearance is not a result of cherry-picking on the history of science.

6 Selective Realism and Meta-Empirical Evidence

6.1 The Selective Realist Response

Both Stanford (2006) and Frost-Arnold (2019) formulate stand-alone arguments against realism, so why bother with putting their observations in the service of holism instead? A well-known response to Stanford's argument comes from *selective* forms of realism:

> substantive theoretical claims that featured in past theories and played a key role in their successes (especially novel predictions) have been incorporated (perhaps somewhat re-interpreted) in subsequent theories and continue to play an important role in making them empirically successful. (Psillos, 2009, 366)

Hence, pace Stanford, it seems that scientists *have* been able to anticipate future-proof, long-standing developments, *at least in part*. Similarly, against Frost-Arnold, selective realists might argue that evidence is not entirely misleading, as it *does* determine at least *some* posits that need to be embraced and retained throughout theory change to support success. As I will show below, a rejoinder to this line of argument is possible by appeal to LCH that, so far as I can see, is unavailable to Stanford and Frost-Arnold.

To make things more concrete, consider how *structural* realists, such as Worrall (1989) or Poincaré (1907), have defended the prototype of SSR (see also Vickers, 2019, 574) by arguing that the recurrence of certain theoretical (in this case: structural) posits conveys confirmation on them which exceeds that received by the remaining posits. Since Maxwell's equations themselves are an outstanding example of retained mathematical structure, let us draw on this example once more to illustrate the subtleties involved in this response to the challenge developed above.

Presumably, we would like to take the fact that the bare Maxwell equations, $\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \wedge \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0$, formally survive the switch from CED to QED to confirm them beyond the degree to which we believe either c or its quantum-replacement q. But it is not so clear just how to conceive of such an incremental confirmation. As I shall argue, the only viable route for conceiving of this confirmatory process appeals to *the very fact that the posit has been retained itself*, and thus to a piece of *meta-empirical* evidence.

Let us first consider a proposal that is consistent with ideas contemplated by Peter Vickers (2017). This proposal suggests to weaken the relevant posit so that it is compatible with the novel evidence as well. Essentially, this means using the observation that the structure represented by the bare equations has resurfaced in different theories as a mere motivation to try and subject the posit itself *directly* to confirmation by the relevant total evidence.[17] However, in the case of the Maxwell equations, it is not so clear just what that means, as $\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \wedge \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0$ is just a schema, not a posit.

In a first attempt, we could try to make this idea more precise by extending the range of the existential quantification to also include quantum operators. But in light of the arguments given above, maybe we should be even more careful and not commit ourselves hastily to a quantum-representation, q, of the fields either. Instead, we might endorse a somewhat open-ended set of representations to anticipate the possibility of rather drastic future changes.[18]

However, the worry here is that this kind of weakening is tantamount to an *immunization strategy* (Boge, 2021, §3.5; Boge, 2020, 4356). For, at any point in time, when we see reason to abandon all the mathematical representations considered so far, the set of allowed representations may simply be enlarged. Relatedly,

[17]This means bypassing the arguments of the previous sections rather than responding to them, but since I think the strategy is futile, this is a matter of indifference.

[18]Thus, let Q denote the set of quantum operators with relevant structure; mathematical objects with co- and contravariant indices that take in pairs $(x,\ |\psi\rangle)$ from the Cartesian product of the relevant space-time manifold $\mathcal{M}$ with an everywhere dense linear submanifold of a Hilbert space, $\mathcal{H}$, and give back states $|\psi'\rangle \in \mathcal{H}$ (Bogolubov et al., 1989, 80). Let also $G \supset Q \cup_{n,m} \Gamma^1_{\langle n,m\rangle}$. We may, furthermore, include an enlarged set of space-times in the definition of G and the domain of x, so that other metrics and an enlarged set, $\mathbb{M}$, of manifolds can be acknowledged (including those that are only locally diffeomorphic to $\mathbb{R}^4$, or those with higher spatial dimensions). Additionally, the tensor-indices might run over a larger subset of the natural numbers, $I \supset \{1, \ldots, 4\}$. And we should also allow operators in front of $f^{\mu\nu}$ that only in some limit behave like four-gradients. And so forth. Thus, we may take it that what has been confirmed by the joint evidence for CED and QED is something like $\exists f \exists j \exists m \forall x \forall \mu, \nu, \gamma, \delta (f \in G \wedge j \in G \wedge (m \in \mathbb{M} \wedge x \in m \wedge \mu, \nu, \gamma, \delta \in I \rightarrow (\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \wedge \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0))$.

notable selective realists have pointed to the posits that are 'essential' for getting the evidence right (Psillos, 1999; Kitcher, 1993; Saatsi, 2005; Chakravartty, 2007; Vickers, 2017), in part in order to calibrate the level of generality of the relevant posits (especially Vickers, 2017, 3226). But what 'essential' means here is all but easily spelled out: It might be possible to define a posit as essential in *relative* terms if it is the weakest posit required for getting the evidence right over some class of theories that successively save increasing amounts of evidence, while standing in a suitable continuity to one another (Boge, 2021, 18). In contrast, it seems like a formidable task to define the essentiality of posits in *absolute* terms (Vickers, 2013, 207 ff.).

An alternative attempt might be to try and remove all the apparent ontological surplus:

$$\exists f \exists j \forall x \forall \mu, \nu, \gamma, \delta (\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \land \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0)). \quad \text{(Maxwell 3)}$$

Of course, this is not quite ontologically neutral yet, as we have indirectly restricted the scope of f and j by the tensor-indices. But more importantly, we already cannot get from this somewhat deprecated version of the Maxwell-equations to any predictions that could be tested, as should be clear from Section 3: It is only by equipping f and j (and even x) with more structure that any calculations can get off the ground, or more concrete models can be defined that spell out the consequences of the equations for the concrete situation at hand. The mere assertion of '*some* multi-component functions of *something*' behaving in certain ways is clearly insufficient for this: (Maxwell 3) does not even unequivocally speak about goings-on in space-time. How could we rationally alter the credence for a given evidence-claim based on this deprecated assumption? The burden is on the realist to show how this might work.

6.2 The Need for Meta-Empirical Evidence

According to the foregoing, the direct route, which takes the fact that a given posit has been successful in different theories or hypotheses as a mere *motivation* for subjecting *it* to confirmation by empirical evidence, does not seem workable: Without committing ourselves to more concrete claims about what electromagnetic fields or tectonic plates *are* and *do*, we cannot even begin to estimate their impact on potential observations and experiments.

The alternative might be to take the very fact that the formula $\partial_\mu f^{\mu\nu}(x) \propto j^\nu(x) \wedge \varepsilon^{\mu\nu\gamma\delta}\partial_\mu f_{\gamma\delta}(x) = 0$ has survived the switch from CED to QED to be itself confirmatory. But *ipso facto*, this means accepting meta-empirical facts as relevant for confirmation: That the formula resurfaces throughout theory change – or that non-rigid plates are still considered 'tectonic' and to move – is a fact about the *context of the inquiry*, not about any measurements or observations that could be cited in support of Maxwell's equations or plate tectonics.

I am not at all in the business here of defending meta-empirical confirmation, nor of offering a concrete model as to how it might work in this particular case. In fact, in the next section, I will offer at least some reasons to vouch for the anti-realist conclusion that holism severely limits our ability to confirm substantive hypotheses instead. All I am saying here is that the *only* piece of evidence the SSRist could advance as confirmatory of some posit she might want to commit herself to, when this posit is detached from the encompassing hypothesis, is a piece of meta-empirical evidence.

Consider a reasonable attempt to avoid this conclusion, that meta-empirical evidence must be invoked, which might proceed by slightly re-framing the problem. For example, one might instead see it as a *sufficient condition for holistic top-down confirmation* that a given posit has been contained in two distinct hypotheses and both of these have received confirmation from evidence – much like the deprecated Maxwell-equations in relation to CED ad QED. One might then try to botch together the collective evidence in support of the posit, based on the fact that it was evidence for both (or several) more encompassing, rival hypotheses. More formally:

(4.) If posit ℓ is an evidence-transcending content-element of both h_1 and h_2, then $P(\ell|e_1) > P(\ell)$ and $P(\ell|e_2) > P(\ell)$, where e_1 and e_2 confirm h_1 and h_2, respectively.

This is no obvious sense an invocation of meta-empirical confirmation. But a surprising result, proven in the Appendix (A 3), is that under fairly reasonable assumptions, $P(\ell|a) > P(\ell)$, where a is any instantiation of the antecedent of (4.). That is, under these assumptions, (4.) *ipso facto* yields an instance of meta-empirical confirmation: The very fact that the posit, ℓ, resurfaces in two distinct, rival hypotheses has a confirmatory impact, on a standard Bayesian account.

One assumption needed to prove this is that (4.) itself does not make it any more likely that the posit, ℓ, is true and confirmed by the evidences e_i than that the posit is true and a content part of both confirmed hypotheses.[19] The second is an unproblematic application of van Fraassen's (1984) special reflection principle, which is 'unproblematic' insofar as it is, by construction, not subject to pertinent counter-examples (e.g. Lin, 2024, §3.6). Since these assumptions seem rather innocent, I thus maintain that SSR can *only* respond to the holistic challenge by calling upon meta-empirical evidence.

6.3 A Dilemma

If everything I have said so far is correct, the realist thus faces a dilemma: Either, she must abandon the approach, thus succumbing to the anti-realist challenge, or embrace meta-empirical evidence. Dawid (2013) and Dawid et al. (2015) have prominently argued that certain pieces of meta-empirical evidence can, indeed, confirm relevant hypotheses. The evidence in question concerns the fact "that scientists have not found any alternatives to a specific solution of a research problem" (Dawid et al., 2015, 215), and so just as well an observation made on the context of inquiry (Dawid, 2022). As is well known, Dawid et al. (2015) devise a Bayesian argument they call the 'No Alternatives Argument', to the effect that the only discovered alternative is incrementally confirmed by the meta-empirical fact.

No-alternatives style reasoning has been considered controversial in delivering a direct argument in support of realism (see Chall, 2018; Dawid, 2021). Dawid et al. (2015, 216) actually state that they "are only interested in arriving at empirically adequate theories and not in the more ambitious goal of finding theories that are true under a given interpretation", but they also "conjecture that the validity of [inference to the best explanation] can sometimes be analysed in terms of a [No Alternatives Argument]". However, since a very specific kind of meta-empirical evidence is here argued to deliver the major support for a very specific form of realism, we should not get tangled up with this debate: The basic issue to be addressed here is whether we should allow meta-empirical evidence to have confirmational impact *at all*.

Obviously, some powerful intuitions speak against this:

[19] Assuming otherwise would essentially amount to a petitio, as it would mean that, upon acceptance of the principle, one considers it more likely for the select posit to become confirmed by the multiple pieces of evidence than that two confirmed hypotheses have shared content at all. No selective realist should want to commit herself to this.

> We create theories with our intelligence, use non-empirical arguments to grow confidence in them, but then *ask nature* if they are right or wrong. They are often wrong. (Rovelli, 2019, 122, orig. emph.)

Of course, it is beyond the scope of this paper to offer a detailed criticism of meta-empirical confirmation. Suffice it so say that there are a number of different criticisms, such as that a significant degree of confirmation is only possible via gerrymandering (Menon, 2019; although see Dawid, 2025), or that it can only work under unrealistic assumptions (Oriti, 2019; although see Dawid, 2022). It seems equally clear, though, that meta-empirical evidence is *somehow* epistemically important. For instance, a whole approach to Hume's problem of induction revolves around strategies for observing how different predictors perform and then following them in sophisticated ways – and thus around exploiting evidence on the meta-level (Schurz, 2019; also Douven, 2023; Sterkenburg, 2021). Yet, even if we side with the predictions of a seeming successful clairvoyant, say, we need not assume that she *is* indeed a clairvoyant: Maybe she just has an unlikely long lucky streak. Thus, taking the fact that meta-empirical evidence is epistemically important to imply that it must be in the same sense truth-conducive as *empirical* evidence is a *non sequitur*.

So, if we have reservations about meta-empirical *confirmation*, what else might meta-empirical evidence tell us about the status of a given hypothesis? A particularly insightful clarification has been offered by Cabrera (2021). Cabrera argues that there is a conflation between *justification* and *pursuit* going on in evaluations of the No Alternatives-Argument: The meta-empirical pieces of evidence invoked by Dawid et al. (2015) may increase our confidence in the pursuitworthyness of the relevant hypothesis; but should they thereby count as justifying belief in it, in the same sense as successfully predicted experimental outcomes or field observations may?

The possibility of applying this distinction is connected to a notorious feature of Bayesian models: That they can be used to model various types of attitude changes; including, e.g., those connected to abductive (Romeijn, 2013) or analogical reasoning (Dardashti et al., 2019; Crowther et al., 2021). Hence, a careful examination of what these changes amount to is always indicated. Note that this is consistent with the treatment of Bayesianism I had advocated in Sect. 3: If there is confirmation, we can model this as a probability-update, but that doesn't mean that every probability-update is an instance of confirmation.

The distinction between justification and pursuit invoked by Cabrera is paralleled by a distinction between belief and acceptance (Frost-Arnold, 2014; see similarly Smeenk, 2019) which, in turn, goes back to Bratman (1992) and has a precursor in (van Fraassen, 1980). Bratman (1992, 3–9) characterizes belief as context independent, shaped primarily by evidence, involuntary, and ideally subjectable to integration. Acceptance, in contrast, diverges from belief on some or all of these features. Crucially, Bratman considers acceptance to be a presupposition of practical reasoning; that is, rational decision-making in contexts that require *action*.

However, acceptance thus construed is an attitude perfectly consistent with anti-realism:

> [...] the sophisticated instrumentalist recognizes that our leading scientific theories are the best foundation and starting point we have not only for uncovering new and unexpected phenomena, but also for opening up new areas and paths of inquiry, and in guiding ourselves to the even more powerful conceptions of natural domains that will ultimately replace the ones we now have. (Stanford, 2006, 207)

> [...] it is possible for an instrumentalist to defend the view that some 'theoretical assertions' are instrumental devices—to promote understanding of relationships between phenomena [...]—and even that such devices are indispensable for good science. (Rowbottom, 2019, 32)

> where can we turn for reliable facts, insightful theories, and guidelines for action? [...] science and the scientific attitude constitute our best hope. But we are often distracted by an impossible ideal of scientific knowledge as proven universal truth about some ultimate reality. (Chang, 2022, 1)

Given the problematic status of meta-empirical confirmation, I thus suggest to take the anti-realist horn of the dilemma: It is preferable to view the situation in terms of acceptance rather than confirmation, so we are entitled to accept the Maxwell equations because they continue to be part of distinct, rival hypotheses that, on the whole, make successful predictions. But that fact itself does not confirm the Maxwell equations, in the sense of promoting their truth. Following the arguments presented from Sect. 1 onward, this means that, unless we embrace a meta-empirical route to confirmation, LCH supported by UP does create a challenge for realism that, unlike the problems of unconceived alternatives and misleading evidence, cannot be avoided by moving to selective forms of realism.

7 Conclusions

I have argued that local confirmational holism (LCH) and underdetermination can be coherently combined into a challenge for realism: If holism is interpreted to say that every logico-semantic part of a hypothesis depends, for its confirmation, on the confirmation of the hypothesis as a whole, we are forced to formulate conditions that determine whether the confirmation received by the whole can be transferred to a given part. Underdetermination suggests that certain plausible necessary conditions for this are typically not met, so the confirmation received by the whole cannot simply be transferred to the parts. But if this is true, the confirmation received by the whole remains bounded by our prior credences for the parts.

The obvious response I addressed is the response from selective realism: If a posit is preserved throughout theory change, or a generalization of it occurs in successor theories, then this posit (or its generalization) *is* confirmed, even beyond the degree to which the containing hypothesis is confirmed. However, LCH also implies that we cannot directly confirm the posit by the collective evidence. Hence, only the meta-empirical fact that the posit is retained could confirm it beyond the degree to which any containing hypothesis is confirmed. And if one simply makes it a *sufficient condition* for transferring confirmation down from distinct encompassing hypotheses to their joint parts that a posit *is* part of such distinct, evidentially supported hypotheses, this provably yields an instance of meta-empirical confirmation as well. Hence, the realist is forced into a dilemma: Either she rejects even selective forms of realism, or she must embrace confirmation by meta-empirical evidence, i.e., the very fact that the posit did survive theory-change. As I have argued, the controversial status of meta-empirical confirmation makes the anti-realist horn look rather attractive in comparison.

Appendix

A 1. *Let h entail c. Then there are probabilistic credence functions P so that $P(h|e) > P(e)$ but $P(c|e) \leq P(c)$ for some e.*

Proof. Assume that $P(c) = .4 > P(h), P(e) = .5, P(h \wedge e) = P(c \wedge e) = .2$. Then $P(h|e) = .4 > P(h)$, by assumption, but also $P(c|e) = .4 = P(c)$. □

A 2. *For any given h and any given piece of evidence e_i, $h \vee \neg e_i$ is an implication of h that is provably not confirmed by e_i.*

Proof. $P(h \vee \neg e_i|e_i) = P((h \vee \neg e_i) \wedge e_i)/P(e_i) = [P(h \wedge e_i) + P(e_i \wedge \neg e_i)]/P(e_i) = P(h|e_i)$. Hence, assume $P(h|e_i) > P(h \vee \neg e_i)$. Then $P(h \wedge e_i) > P(e_i)P(h \vee \neg e_i)$, which is false, since $0 \leq P(h \vee \neg e_i) \leq 1$ and $P(h \wedge e_i) \leq P(e_i)$. □

A 3. *Let ℓ an evidence transcending content-part of two distinct hypotheses h_1 and h_2, a an arbitrary instantiation of (4.)'s antecedent, and c_i a corresponding instantiation of one of the conjuncts of (4.)'s consequent. Further, assume that an agent accepts (4.) and is subject to the special reflection principle $\mathrm{Pr}(h|\mathrm{Pr}^{t^+}(h) = r) = r$ (Pr a generic probabilistic credence function) at least under the condition that all that changes until future time t^+ is that some new piece of evidence is obtained. Then if $P(\ell \wedge c_i) \leq P(\ell \wedge a)$, where P is implicitly conditioned on (4.), we obtain $P(\ell|a) > P(a)$.*

Proof. We formalize (4.) as $\theta \equiv \forall \ell, e_1, e_2, h_1, h_2(\ell \in C_{\oplus e_1}(h_1) \cap C_{\oplus e_2}(h_2) \wedge \mathrm{Pr}(h_1|e_1) > \mathrm{Pr}(h_1) \wedge \mathrm{Pr}(h_2|e_2) > \mathrm{Pr}(h_2) \rightarrow \mathrm{Pr}(\ell|e_1) > \mathrm{Pr}(\ell) \wedge \mathrm{Pr}(\ell|e_2) > \mathrm{Pr}(\ell))$, where $C_{\oplus e_j}(h_i)$ are the content-parts of h_i transcending evidence e_j, and Pr is a generic credence function that may itself be conditioned on θ in application. Now assume that one of the pieces of evidence, e_i, is in fact learned at t^+ just slightly after a, and that nothing else changes. Accordingly, we define $P^{t^+}(\cdot) := P(\cdot|e_i)$. Further, let $P(\ell|e_i) =: \delta > 0$, which, by assumption, is greater than $P(\ell)$, and let c_i be the relevant instantiation of $P(\ell|e_i) = \delta > P(\ell)$. Then, special reflection yields $P(\ell|c_i) = P(\ell|P^{t^+}(\ell) = \delta) = \delta$. Since $a \wedge \theta$ entails $c_i \wedge \theta$, we have $p(a \wedge \theta) \leq p(c_i \wedge \theta)$, where p denotes the agent's credences not conditioned on θ. Thus, $P(\ell|a) = \frac{p(\ell \wedge a \wedge \theta)}{p(a \wedge \theta)} \geq \frac{p(\ell \wedge a \wedge \theta)}{p(c_i \wedge \theta)}$. Furthermore, since $\ell \wedge a \wedge \theta$ equally entails $\ell \wedge c_i \wedge \theta$, we also have $p(\ell \wedge a \wedge \theta) \leq p(\ell \wedge c_i \wedge \theta)$. However, by assumption, $P(\ell \wedge c_i) \leq P(\ell \wedge a)$ and so $p(\ell \wedge c_i \wedge \theta) \leq p(\ell \wedge a \wedge \theta)$, whence $p(\ell \wedge c_i \wedge \theta) = p(\ell \wedge a \wedge \theta)$. Hence, $P(\ell|a) \geq \frac{p(\ell \wedge a \wedge \theta)}{p(c_i \wedge \theta)} = \frac{p(\ell \wedge c_i \wedge \theta)}{p(c_i \wedge \theta)} = P(\ell|c_i) = \delta > P(\ell)$. □

Acknowledgments I have profited from comments by Nico Formanek, James Fraser, Gerhard Schurz, and in general the audience of the *Philosophy of Science Mini-Workshop* held October 2022 at Wuppertal University. I have also profited from comments by a number of anonymous referees at different journals, particularly the two referees for this journal. The main research for this paper was done during my employment with the research unit *The Epistemology of the Large Hadron Collider*, funded by the German Research Foundation (DFG; grant FOR 2063). Several revisions were made during my leadership of the Emmy Noether group *UDNN: Scientific Understanding and Deep Neural Networks*, also funded by the DFG (project number 508844757). I hence acknowledge the funding.

Declaration of interest: None.

Declaration of generative AI in scientific writing: None.

References

Boge, F. J. (2020). An argument against global no miracles arguments. *Synthese*, 197(10):4341–4363.

Boge, F. J. (2021). Incompatibility and the pessimistic induction: a challenge for selective realism. *European Journal for Philosophy of Science*, 11(2):1–31.

Boge, F. J. (2023). The positive argument against scientific realism. *Journal for General Philosophy of Science*, 54(4):535–566.

Bogolubov, N., Logunov, A., Oksak, A., and Todorov, I. (1989). *General Principles of Quantum Field Theory*. Springer Netherlands.

Bolstad, W. and Curran, J. (2017). *Introduction to Bayesian Statistics*. Wiley, third edition.

Boyd, R. N. (1973). Realism, underdetermination, and a causal theory of evidence. *Nous*, pages 1–12.

Bratman, M. E. (1992). Practical reasoning and acceptance in a context. *Mind*, 101(401):1–15.

Brimacombe, M. (2018). *Likelihood Methods in Biology and Ecology: A Modern Approach to Statistics*. CRC Press.

Cabrera, F. (2021). String theory, non-empirical theory assessment, and the context of pursuit. *Synthese*, 198:3671–3699.

Carnap, R. (2001 [1937]). *Logical Syntax of Language*. London: Routledge.

Cartwright, N. (1983). *How the laws of physics lie*. OUP Oxford.

Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge: Cambridge University Press.

Chall, C. (2018). Doubts for dawid's non-empirical theory assessment. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 63:128–135.

Chang, H. (2022). *Realism for Realistic People: A New Pragmatist Philosophy of Science*. Cambridge University Press.

Crowther, K., Linnemann, N. S., and Wüthrich, C. (2021). What we cannot learn from analogue experiments. *Synthese*, 198(16):3701–3726.

Dardashti, R., Hartmann, S., Thébault, K., and Winsberg, E. (2019). Hawking radiation and analogue experiments: A bayesian analysis. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 67:1–11.

Das, S. and Dürr, D. (2019). Arrival time distributions of spin-1/2 particles. *Scientific Reports*, 9(2242). https://doi.org/10.1038/s41598-018-38261-4.

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

Dawid, R. (2021). The role of meta-empirical theory assessment in the acceptance of atomism. *Studies in History and Philosophy of Science Part A*, 90:50–60.

Dawid, R. (2022). Meta-empirical confirmation: Addressing three points of criticism. *Studies in History and Philosophy of Science*, 93:66–71.

Dawid, R. (2025). Does the no alternatives argument need gerrymandering to be significant? *The British Journal for the Philosophy of Science*, 76(1):209–222.

Dawid, R., Hartmann, S., and Sprenger, J. (2015). The no alternatives argument. *The British Journal for the Philosophy of Science*, 66(1):213–234.

Douglas, H. E. (2009). Reintroducing prediction to explanation. *Philosophy of Science*, 76(4):444–463.

Douven, I. (2023). Explaining the success of induction. *The British Journal for the Philosophy of Science*, 74(2):381–404.

Duhem, P. (1954[1914]). *The Aim and Structure of Physical Theory*. Princeton University Press, second edition. translated by Philip P. Wiener.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4.

Fahrbach, L. (2021). We think, they thought: A critique of the pessimistic meta-meta induction. In Lyons, T. D. and Vickers, P., editors, *Contemporary Scientific Realism: The Challenge From the History of Science*, pages 283–310. Oxford University Press.

Fallis, D. and Lewis, P. (2014). Misleading evidence. *SSRN*. https://ssrn.com/abstract=2462771orhttp://dx.doi.org/10.2139/ssrn.2462771.

Fodor, J. and Lepore, E. (1992). *Holism: A Shopper's Guide*. Oxford, Cambridge (MA): Blackwell.

Frank, S. A. (2012). Natural selection. iv. the price equation. *Journal of evolutionary biology*, 25(6):1002–1019.

Fraser, J. D. and Vickers, P. (2022). Knowledge of the quantum domain: An overlap strategy. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/721635.

Frost-Arnold, G. (2019). How to Be a Historically Motivated Antirealist: The Problem of Misleading Evidence. *Philosophy of Science*, 86(5):906–917.

Frost-Arnold, K. (2014). The cognitive attitude of rational trust. *Synthese*, 191(9):1957–1974.

Gardner, A. (2020). Price's equation made clear. *Philosophical Transactions of the Royal Society B*, 375(1797):20190361.

Giere, R. (2010). *Explaining Science: A Cognitive Approach*. University of Chicago Press.

Gillies, D. (2000). *Philosophical Theories of Probability*. London, New York: Routledge.

Gould, S. J. and Lewontin, R. C. (1979). The spandrels of san marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161):581–598.

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge university press.

Keith, M. (2001). Evidence for a plate tectonics debate. *Earth-Science Reviews*, 55(3):235–336.

Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford: Oxford University Press.

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

Laudan, L. and Leplin, J. (1991). Empirical equivalence and underdetermination. *The journal of philosophy*, 88(9):449–472.

Lawrence, A. (2019). Probability in physics. *Undergraduate Lecture Notes in Physics, Springer-Verlag, Cham, https://doi. org/10.1007/978-3-030-04544-9*.

Lin, H. (2024). Bayesian Epistemology. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2024 edition.

Luque, V. J. and Baravalle, L. (2021). The mirror of physics: on how the price equation can unify evolutionary biology. *Synthese*. https://doi.org/10.1007/s11229-021-03339-6.

Magnus, P. (2010). Inductions, red herrings, and the best explanation for the mixed record of science. *The British Journal for the Philosophy of Science*, 61(4):803–819.

Magnus, P. (2013). Philosophy of science in the twenty-first century. *Metaphilosophy*, 44(1-2):48–52.

Manson, R. (2003). Stripes on the sea floor. In Oreskes, N., editor, *Plate tectonics*, pages 31–45. Westview Press.

Menon, T. (2019). On the viability of the no alternatives argument. *Studies in History and Philosophy of Science Part A*, 76:69–75.

Okasha, S. (2002). Underdetermination, holism and the theory/data distinction. *The Philosophical Quarterly*, 52(208):303–319.

Okasha, S. and Otsuka, J. (2020). The price equation and the causal analysis of evolutionary change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1797):20190365.

Oreskes, N. (2003). From continental drift to plate tectonics. In Oreskes, N., editor, *Plate tectonics*, pages 3–27. Westview Press.

Oriti, D. (2019). No alternative to proliferation. In Dardashti, R., Dawid, R., and Thebault, K., editors, *Why Trust a Theory? Epistemology of Fundamental Physics*. Cambridge University Press.

Orzack, S. H. and Sober, E. (1994). Optimality models and the test of adaptationism. *The American Naturalist*, 143(3):361–380.

Poincaré, H. (1907). *The Value of Science*. The Science Press.

Popper, K. and Miller, D. (1983). A proof of the impossibility of inductive probability. *Nature*, 302(5910):687–688.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. London; New York: Routledge.

Psillos, S. (2009). "grasping at realist straws." review of *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*, by p. kyle stanford. *Metascience*, 18(3):363–370.

Quine, W. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, 60(1):20–43.

Ribeiro, A. and Mateus, A. (2012). *Soft Plate and Impact Tectonics*. Springer Berlin Heidelberg.

Ritson, S. and Staley, K. (2021). How uncertainty can save measurement from circularity and holism. *Studies in History and Philosophy of Science Part A*, 85:155–165.

Roche, W. and Sober, E. (2017). Explanation= unification? a new criticism of friedman's theory and a reply to an old one. *Philosophy of Science*, 84(3):391–413.

Romeijn, J.-W. (2013). Abducted by bayesians? *Journal of Applied Logic*, 11(4):430–439. Combining Probability and Logic: Papers from Progic 2011.

Rovelli, C. (2019). The dangers of non-empirical confirmation. In Dardashti, R., Dawid, R., and Thébault, K., editors, *Why Trust a Theory?: Epistemology of Fundamental Physics*, page 120–124. Cambridge University Press.

Rowbottom, D. (2011). *Popper's Critical Rationalism: A Philosophical Investigation*. Taylor & Francis.

Rowbottom, D. (2019). *The Instrument of Science: Scientific Anti-Realism Revitalised*. Taylor & Francis.

Saatsi, J. (2005). Reconsidering the fresnel–maxwell case study. *Studies in History and Philosophy of Science*, 36:509–538.

Salmon, W. (1990). Rationality and objectivity in science or tom kuhn meets tom bayes. In Savage, C. W., editor, *Scientific Theories*, pages 175–204. University of Minnesota Press.

Schaeben, H., Kroner, U., and Stephan, T. (2024). Mathematical fundamentals of spherical kinematics of plate tectonics in terms of quaternions. *Mathematical Methods in the Applied Sciences*, 47(6):4469–4496.

Schippers, M. and Schurz, G. (2020). Genuine confirmation and tacking by conjunction. *The British Journal for the Philosophy of Science*, 71(1):321–352.

Schurz, G. (2014). Bayesian pseudo-confirmation, use-novelty, and genuine confirmation. *Studies in History* *and Philosophy of Science Part A*, 45:87–96.

Schurz, G. (2019). *Hume's problem solved: The optimality of meta-induction*. Mit Press.

Schurz, G. (2022). Tacking by conjunction, genuine confirmation and convergence to certainty. *European Journal for Philosophy of Science*, 12(3):1–18.

Schurz, G. and Weingartner, P. (2010). Zwart and Franssen's impossibility theorem holds for possible-world-accounts but not for consequence-accounts to verisimilitude. *Synthese*, 172(3):415–436.

Sklar, L. (1975). Methodological conservatism. *The Philosophical Review*, 84(3):374–400.

Smeenk, C. (2019). Gaining access to the early universe. In Dardashti, R., Dawid, R., and Thébault, K., editors, *Why Trust a Theory? Epistemology of Fundamental Physics*, pages 315–335. Cambridge University Press.

Sprenger, J. (2011). Hypothetico-deductive confirmation. *Philosophy Compass*, 6(7):497–508.

Stanford, K. (2023). Underdetermination of Scientific Theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2021 edition.

Stanford, P. K. (2001). Refusing the devil's bargain: What kind of underdetermination should we take seriously? *Philosophy of Science*, 68(S3):S1–S12.

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

Sterkenburg, T. F. (2021). On explaining the success of induction. *The British Journal for the Philosophy of Science*. https://doi.org/10.1086/717068.

van Fraassen, B. C. (1980). *The scientific image*. Oxford University Press.

van Fraassen, C. (1984). Belief and the will. *The Journal of Philosophy*, 81(5):235–256.

Vickers, P. (2013). A confrontation of convergent realism. *Philosophy of Science*, 80(2):189–211.

Vickers, P. (2017). Understanding the selective realist defence against the PMI. *Synthese*, 194:3221–3232.

Vickers, P. (2019). Towards a realistic success-to-truth inference for scientific realism. *Synthese*, 196:571–585.

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

Wray, K. (2018). *Resisting Scientific Realism*. Cambridge University Press.