9
The Historical Challenge to Realism and Essential Deployment

Mario Alai

Department of Pure and Applied Sciences—Università di Urbino Carlo Bo
Académie Internationale de Philosophie des Sciences
mario.alai@uniurb.it

9.1 Deployment Realism

Scientific realists use the "No Miracle Argument" (**NMA**): it would be a miracle if theories were false, yet got right so many novel and risky predictions. Hence, predictively successful theories are true. Of course, one could easily make up a theory *with completely false theoretical assumptions which predicted a phenomenon* ***P*** (call it a $\mathbf{F^P}$-theory) if she knew **P** in advance and used it in framing the theory. But how could she think of a $\mathbf{F^P}$-theory, without knowing **P**? Or knowing **P** but without using it in building the theory? In fact, it is puzzling how one could have built a $\mathbf{F^P}$-theory even if she used **P** *inessentially*: suppose Jill built a $\mathbf{F^P}$-theory by knowing and using P, but she *could* have done without it, because, quite independently of her, John built *the same* theory without using **P**. This I call Jill using **P** *inessentially*, and it is something hard to explain, because it is understandable how the theory was built by Jill, but not by John (Lipton 1991, 166; Alai 2014c, 301).

So, how could one find a $\mathbf{F^P}$-theory without using **P** essentially? Well, one could do this if **P** were very poor in content, i.e., very probable: any false theory could predict the existence of *some* new planet *somewhere* in the universe. But it took Newton's theory to predict *precisely* the existence of a planet of *such and such* mass with *such and such* orbit as Neptune. So, the riskier a prediction is (i.e., the more precise, rich in content, hence improbable it is), the less likely it is that it has been made by a completely false theory.

It might be objected that since a falsity may entail a truth, even false theories might make novel predictions. Granted, in the Hyperuranion of possible theories there are some with completely false theoretical assumptions which

Mario Alai, *The Historical Challenge to Realism and Essential Deployment* In: *Contemporary Scientific Realism.* Edited by: Timothy D. Lyons and Peter Vickers, Oxford University Press. © Oxford University Press 2021.
DOI: 10.1093/oso/9780190946814.003.0009

(together with appropriate auxiliary assumptions) make true *novel and improbable* predictions (call them **F**$^{\mathbf{NIP}}$-theories). However, their rate is very small, for it is a logical fact that the less probable a consequence is, the fewer the premises entailing it. Hence, it is extremely improbable that **F**$^{\mathbf{NIP}}$-theories are found *by chance*. But since scientists look for *true* theories, only by chance can they pick a **F**$^{\mathbf{NIP}}$-theory. Therefore, it is extremely unlikely that scientists find a **F**$^{\mathbf{NIP}}$-theory.[1]

For instance, quantum electrodynamics predicted the magnetic moment of the electron to be $1159652359 \times 10^{-12}$, while experiments found $1159652410 \times 10^{-12}$: hence John Wright (2002, 143–144) figured that the probability to get such an accuracy *by chance*, i.e., through a false theory, is as low as 5×10^{-8}.

Here (following Alai 2014c) I will speak of "novel predictions" (**NP**) when the predicted phenomenon is either unknown (historically novel) or not used (use-novel) or not used essentially (functionally novel) by the theorist. So, it is *practically* impossible that a *novel* prediction which in addition is risky, i.e., a highly improbable, was derived from a completely false theory.

Against the claim that success warrants truth, Laudan (1981) had objected that many false theories in the history of science were successful, so realists replied that however those theories didn't have *novel* predictive success.[2] Unfortunately, it turns out that also many false theories had novel success (see the list in Lyons 2002, 70–72). Selective realists acknowledged the problem and argued that the **NMA** does not warrant the *complete* truth of theories, but only their *partial* truth. In particular, for Kitcher's (1993) and Psillos's (1999) "Deployment Realism" (**DR**), we should be committed only to those *particular hypotheses* which were *deployed* in deriving novel predictions: this is Psillos's "*divide et impera*" move against Laudan's "reductio" (Psillos 1999, 102, 108, passim).

A rejoinder came from Timothy Lyons: first, he explained that Laudan's original objection was not "an articulation of the pessimistic *meta-induction* toward the conclusion that *our present theories are probably false*," but a *modus tollens* refutation of the **NMA** (a "meta-*modus tollens*") (Lyons 2002, 64–65; 2006, 557). Secondly, he criticized **DR** by listing a number of false *hypotheses* which were actually deployed in the derivation of various novel risky predictions—including Neptune's discovery (2002; 79–83; 2006). Deployment realists, however, can answer this new challenge by arguing that the **NMA** does not warrant the truth of all hypotheses *actually used* in deriving novel predictions, but only of those which were *essential*, i.e., strictly necessary to those predictions. False hypotheses are no counterexample to the **NMA** if they were used *de facto*, but actually superfluous, i.e., *non-essential*.

[1] Alai 2014c, §1; 2014d, §5; 2016, §4. More on this objection at §9.7.1.
[2] Musgrave (1985, 1988); Lipton (1993, 1994); Psillos (1999); Sankey (2001).

Thus, *essentiality* is crucial to two key claims of **DR**: (1) a predicted phenomenon is novel only if it was not used essentially in building the theory which predicts it; and (2) a hypothesis is most probably true only if it was essential in deriving a novel prediction. Here we shall focus on the latter claim.

According to Psillos (1999, 110) a hypothesis **H** is deployed *essentially* in deriving a novel prediction **NP** if

(1) **NP** follows from **H**, together with some other hypotheses **OHs** and auxiliary assumptions **AA**, but not from **OHs** & **AA** alone;
(2) no *other* hypothesis **H*** is *available* which can do the same job as **H**, viz. is
  (a) compatible with **OHs** and **AA**,
  (b) non-*ad hoc*,
  (c) potentially explanatory, and
  (d) together with **OHs** and **AA** predicts **NP**.

Conditions (1) and (2) capture the idea that **H** is essential, i.e., strictly necessary, when (1) **OHs** without **H** could not predict **NP**, and (2) no alternative hypothesis is available which could substitute **H** in the derivation of **NP**.

9.2 Lyons's Criticism of Psillos's Criterion of Essentiality

However Lyons (2006; 2009) criticized this criterion of essentiality, in particular condition **(2)**. First, he argued that it is too vague to be applicable to any historical case (2006, 542). For instance, it is unclear *when* **H*** should not be *available*: when the theory is put forward? when the prediction is derived? when the prediction is confirmed? at some point in the future? Further unclear points are what "potentially explanatory" means; whether each of the elements of **OHs** and **AA** also need to be "essential" by this criterion; whether the replacement hypothesis needs to be consistent with those components of **OHs** and **AA** which are "essential" for other predictions but unnecessary for the prediction of concern; whether **H*** can result in the loss of other confirmed predictions.

Secondly, according to condition **(2)** a competitor **H*** makes **H** inessential only if **H*** is compatible with **OHs** and **AA**, and it is non-ad hoc (i.e., predicts phenomena not used in constructing **H***: Psillos 1999, 106). But Lyons notices that in actual history "most (if not all) well-respected theories" were either *incompatible* with **OHs** and **AA**, or ad hoc in Psillos's sense, or both (Lyons 2009, 149). So the requirement placed by **(2)** on admissible competitors is "too strict to permit even some of the most exemplary scientific theories . . . to qualify as competitors" (150), and this means that **(2)** is too weak (or practically empty), for most (if not all) possible hypotheses **H** will lack admissible competitors. Psillos didn't understand his conditions for essentiality as a definition, but just as a criterion: he writes "**H** indispensably contributes to the generation of **P** if . . . [(**1**) and (**2**) are met]," not "if *and only if* . . . [(**1**) and (**2**) are met]" (1999, 110). Yet, Lyons's criticisms are at least prima facie well-founded, and must be taken seriously.

But Lyons also claims that his criticisms force deployment realists to abandon the essentiality requirement altogether. Besides, he claims that the "crucial" or "fundamental" insight of **DR** is that we should credit "those and only those constituents that were genuinely responsible for, that actually led scientists to, specific predictions" (2006, 543), no matter whether they were indispensable or not. In fact, he argues that Psillos, after introducing his criterion of essentiality (Psillos 1999, 109–110), "never mentions that rigid criterion again"[3] and fails to show that his case studies comply with it (Lyons 2009, 143). On the contrary, according to Lyons, Psillos often talks of commitment to constituents which simply "generated," or "brought about," or were "responsible for," or "genuinely contributed to," or "really fueled" (Psillos 1999, 108–110) the success of the theory (Lyons 2006, 540). Moreover, Lyons says that the same "fundamental insight" is also embraced, in some form or another, by Kitcher, Niiniluoto, Leplin, and Sankey (538). Thus, he holds, **DR** should get rid of the essentiality requirement and "return to the crucial insight" of actual deployment (543).

At this point, however, Lyons argues that many assumptions which in history were *actually* employed and led to novel predictions are *false*. In fact, in various papers (2002, 2006, 2016, 2017) he discusses a number of striking novel predictions, in *each* of which many false assumptions were employed (we shall briefly review some of them later). Hence, he concludes the **NMA** does not work even when restricted to particular hypotheses, and **DR** is false: "If we take seriously the no miracles argument [i.e., if novel predictions derived from false assumptions are a miracle] . . . we have witnessed numerous miracles" in history (2006, 557).

However, in reply, a number of points should be noticed: first, by accepting such purported "miracles," one is left with the problem of explaining how are they possible. If one wonders how scientist S made the *novel risky prediction* **NP** which was then confirmed, a possible answer is that **NP** followed from true assumptions in S's theory, for all consequences of true assumptions are true. But one cannot answer that **NP** followed from some false hypothesis in S's theory, because that would entail that S had luckily assumed a false hypothesis entailing **NP**, and as mentioned, this is extremely improbable.[4] Lyons (2002; 2003) and

[3] Personal communication.

[4] More on this at §9.7.1. One might undercut the **NMA** and **DR** also by rejecting the need or possibility of explaining unlikely events, or of explanations in general. But I cannot discuss that radical objection here.

others proposed alternative explanations, not involving the truth of the deployed hypotheses, but I have argued elsewhere that they all fail (Alai 2012, 88–89; 2014a).

Secondly, as concerns which one is the "fundamental insight" of **DR** that cannot be abandoned vs. what is merely accessory and should be dropped, two questions must be distinguished: (I) the exegetic question whether Psillos understands "deployment" as *merely actual* deployment, or as *essential* deployment; and (II) the theoretical question of how deployment *should* be understood if we are to give **DR** its best currency in order to correctly assess its tenability.

As to (I), when Psillos claims that we should credit constituents that are "responsible" for success (1999, 108–109), Lyons understands "responsibility" as actual deployment, suggesting that this was the original core intuition of **DR** (2006, 540, 543). But he is wrong, for in the next page Psillos specifies that constituents to be credited are those "that made *essential* contributions" to success and that "have an *indispensable* role" in it, and right after this he spells out his conditions (1)–(2) (109–110). So, by responsibility he clearly means *essential* responsibility.

Even apart from this, when an author says two different things, in the exegesis we are allowed to discount one of them only if they are contradictory. But if they are not contradictory we must keep both, and understand how they can be reconciled. In this case, sometimes Psillos talks of responsibility or contribution, or deployment, etc., without further qualifications, sometimes he introduces the essentiality or indispensability qualification. But these two talks are not contradictory, and can be easily reconciled, since the latter is just a specification of the former. In fact, when talking plainly of responsibility or contribution, Psillos never adds "no matter whether essential or not" or the like. Therefore, even his plain talk of responsibility or contribution must be understood as implying "essential" or "indispensable."

Even if in his case studies (1999, ch. 6) Psillos fails to explicitly refer back to his *criterion* of essentiality and to explicitly show how it applies, he does use the essentiality *requirement*. For instance, he argues that we need not be committed to the caloric hypothesis, although it was actually deployed in Laplace's prediction of the speed of sound in air, because that prediction "did not depend on this hypothesis" (121). Hence, he obviously would not give up his essentiality requirement. Moreover, in the absence of any indication to the contrary, and as this quotation suggests, we must assume that he understands the essentiality requirement precisely in terms of his essentiality criterion.

Summing up, from the exegetic point of view, ignoring the essentiality requirement would be misinterpreting Psillos's position. But apart from this, since here we must assess the tenability of **DR**, what matters more is not the exegetical question, but (II) the theoretical question of how **DR** *should* be understood in order to give it the most favorable interpretation. Now again, the answer is the essential deployment interpretation, for it immunizes **DR** from Lyons' purported historical counterexamples.

In fact, to begin with, by Occam's razor essentiality is *required* to the cogency of the **NMA**: in abductions we can assume only what is essential, i.e., the weakest hypothesis sufficient to explain a given effect; but if a hypotheses, although deployed, was not essential in deriving **NP**, it is not essential in explaining its derivation either; therefore deployment realists need not (and must not) be committed to its truth.

Moreover, Psillos's conditions (1)–(2) are just *one* way to spell out essentiality. Therefore, even if Lyons succeeds in criticizing them, he has not shown that no other characterization of essentiality is possible, hence he cannot conclude that this condition *must* be abandoned. In fact, he only *suggests* that it be abandoned, so proposing a version of **DR** deprived of the essentiality requirement which he (wrongly, I argued) believes to be Psillos's "fundamental" version. Therefore, although his refutation of *this* latter version by meta-*modus tollens* from historical cases is effective, it does not affect *actual* **DR**, i.e., the essential deployment version.

Granted, Lyons's analysis of Psillos's conditions (1)–(2) is correct, and his criticisms of (2) are convincing. But I will propose a different condition, which performs the task Psillos had in mind, yet is simpler, less troublesome, and escapes Lyons's criticisms. This condition is not altogether new, having been used or presupposed at various places, I suggest, by Psillos himself (e.g., 1999, 119–121) and explicitly proposed by Vickers, who writes:

> **H** does not merit realist commitment whenever **H** is doing work in the derivation solely in virtue of the fact that it entails some other proposition which itself is sufficient, when combined with the other assumptions in play, for the relevant derivational step (2016, §3).

9.3 An Improved Criterion of Essentiality Escaping Lyons's Criticisms

I propose to substitute the crucial condition **(2)** in Psillos's criterion with:

(2') There is no other hypothesis **H*** which is *proper part* of **H** (hence weaker than **H**) which together with **OHs** and **AA** entails **NP**.

Here in section 9.3 I explain **(2')**, spelling out the intuitive notion of indispensability more precisely through Yablo's (2014) concept of proper parthood, and arguing that the troublesome requirements of Psillos's criterion can be dropped without loss, thus escaping Lyons's criticisms; in section 9.4 I show how my criterion immunizes **DR** from Lyons's historical counterexamples; in section 9.5 I hold that essentiality in this sense cannot be detected prospectively, and Vickers's optimism on the identifiability of non-essentiality should be complemented by some less optimistic considerations; but far from causing problems for realists, this frees them from unreasonable obligations; in section 9.6 I explain in more details why my criterion is enough, although apparently weaker than Psillos's; in section 9.7 I answer some actual and potential objections.

Intuitively, a hypothesis **H*** is a proper part of **H** iff its content is part of the content of **H** but does not exhaust it. At first approximation and for most purposes, a part of **H** is therefore a hypothesis entailed by **H** but not entailing it. But this is not always enough: for instance consider

(i) It's Spring

and

(ii) It's April.

As (i) is entailed by (ii) without entailing it, (i) is a proper part of (ii). But

(iii) It's April *or* the streets are wet

is implied by (ii) without entailing it, however it is not a (proper) part of it, for it brings in a different subject matter. Therefore, just characterizing parthood in terms of entailment will not always do for our purposes. Instead, we must understand parthood as Yablo (2014):[5] **H*** is a part of **H** iff it is entailed by **H** *and preserves its subject matter*. The subject matter of a proposition consists of two classes: that of its truthmakers (or reasons why it is true) and that of its falsemakers (or reasons why it is false). A proposition **H*** entailed by **H** preserves **H**'s subject matter, hence, is *part* of it, iff every truthmaker/falsemaker of **H*** is entailed by a truthmaker/falsemaker of **H**. Moreover, **H*** is a *proper* part of **H** iff every truthmaker/falsemaker of **H*** is entailed by a truthmaker/falsemaker of **H**, *but not vice versa*. This explains the intuitive idea that (iii), although entailed by (ii), is not part of it: because some truthmakers of (iii) (e.g., "it's raining") are not entailed by any truthmaker of (ii).[6]

[5] I thank Matteo Plebani for pointing this out to me. See Plebani (2017).

[6] Although Yablo's approach helps us to properly explain essentiality, it was not designed to this purpose, but to answer the general question of "aboutness," i.e., the subject-matter of sentences: why and how sentences with the same truth-conditions say different things, like "here is a sofa" and "here is the front of a currently existing sofa, and behind it is the back." Thus, it offers a natural solution to a wide variety of problems concerning agreement, orders, possibilities, priority, explanation,

Condition **(2')** expresses precisely the abovementioned Occam's requirement, namely that **H** is not redundant, i.e., that one could not explain the derivation of **NP** by assuming the truth of *something less* than **H**. For instance, suppose that instead of his actual gravitation theory Newton had advanced the following theory:

(**H**) inside each massive body there resides a demi-god, which attracts the demi-gods dwelling in each other body by a force $F = Gm_1m_2/r^2$.

H would predict the same novel phenomena as Newton's actual theory, but we wouldn't need to believe it, because it is redundant, i.e., not *essential* to those predictions. In fact, **H** consists in the following assumptions:

(**H***) each body attracts all other bodies by a force $F = Gm_1m_2/r^2$
(**H**$^{d\text{-}g}$) F is exerted by a demi-god residing inside each body.

H would violate condition **(2')**, because there would be another hypothesis **H*** (i.e., Newton's actual theory) which is a proper part of **H** and sufficient to derive its novel predictions.

This alternative condition **(2')** avoids Lyons's vagueness and emptiness objections to Psillos's condition **(2)**. First, **(2')** is not too vague because of five reasons: (i) there is no question about when the alternative **H*** is *available*, because **(2')** does not exclude just that **H*** is available, but even that it is logically possible (for if it were possible, **H** would be redundant in the prediction of **NP**, hence possibly false); (ii) there is no need to specify what "explanatory" means, since **(2')** has no such requirement; (iii) it is not required that the elements of **OHs** and **AA** are also essential;[7] (iv) since **(2')** excludes only hypotheses **H*** which are part of **H**, ipso facto such hypotheses are compatible with all the components of **OHs** and **AA** with which **H** is compatible; (v) alternative hypotheses **H*** are excluded no matter whether substituting **H** by **H*** results in the loss of other confirmed predictions or not: if **H** allows us to derive two predictions **NP**$_1$ and **NP**$_2$, and the weaker **H*** allows us to derive **NP**$_1$ but not **NP**$_2$, then **H** is not essential with respect to **NP**$_1$, but it may be essential with respect to **NP**$_2$, in which case it is most probably true.

Lyons's second criticism was that **(2)** is practically empty, because it excludes only alternative hypotheses which are compatible with **OHs** and **AA**, while

knowledge, partial truth, and confirmation. For instance, a white marble does not confirm (1) "All crows are black" because (1) does not say the same as (2) "All non-black things are non-crows," and Yablo explains why (2014, pp. xiv, 7–8).

[7] Whether they are or not can be decided in the same way as for **H**. In section 9.6 I shall notice that different but compatible assumptions may all be essential for a given novel prediction.

in most actual cases alternative hypotheses are not compatible with **OHs** and **AA**. My condition **(2')** escapes this problem, since it excludes *all* the alternative hypotheses which are part of **H** and can derive **NP**, without exceptions. In fact, in section 9.4 I shall mention some historical examples of hypotheses which violated condition **(2')**, confirming that it is not empty.

Yet, it might seem that here lurks a further difficult problem: as just said, all the alternative hypotheses excluded by **(2')** are compatible with **OHs** and **AA**, since they are part of **H**. But how about those alternative hypotheses which are *not* compatible with **OHs** and **AA**, because they are *not* part of **H**? Lyons stresses that they exist in many historical cases, and if so, don't they make the original hypothesis **H** superfluous, i.e., non-essential? Why doesn't my criterion exclude such hypotheses?

In other words, Psillos's condition **(2)** for essentiality is of medium strength: it excludes all the hypotheses which can derive **NP** when joined with **OHs** and **AA**. Lyons's criticism is that it is too weak, and his implicit suggestion is that it should be strengthened to exclude even the hypotheses which can derive **NP** in conjunction with different collateral assumptions. My condition **(2')**, instead, is even weaker than Psillos's (for it excludes only a subset of the alternative hypotheses excluded by **(2)**, i.e., only those which are part of **H**), and this is why it escapes Lyons's first criticism: how can it be enough? In particular, shouldn't we require that **H** is *unique*, i.e., that there is no alternative hypothesis **H'** incompatible with **H**, which (no matter whether compatible and joined with **OHs** and **AA**, or incompatible with them and joined with different hypotheses and auxiliary assumptions) can derive **NP**? For if it existed, we couldn't know which of them is true.

I answer that we don't need to explicitly exclude any other **H'** except those which are part of **H** (so, we don't need Psillos's condition **(2)** with its drawbacks, nor its strengthening implicitly suggested by Lyons), because the required uniqueness of **H** is already ensured by condition **(2')** together with the *risky* character of the prediction **NP**. In fact, just as it happens for *theories*, it is practically impossible to find a *completely false* hypothesis making risky novel predictions. A false hypothesis **H** may yield a risky novel prediction **NP**, only if it is *partly* true: i.e., if it has a true part **H*** from which **NP** can be derived. But in this case, the essential role in the derivation of **NP** is played by **H***, **H** being inessential by my condition **(2')**.

This much we know a priori from the **NMA** (as spelled out in the five initial paragraphs of section 9.1), and anti-realists have not been able to spot any inferential fallacy in that argument. However, they challenged it by apparent counterexamples, i.e., by citing historical instances of false hypotheses which were supposedly deployed essentially in deriving novel predictions. But in each of those cases it can be shown that either (a) those hypotheses were not actually false or the predictions not actually true, or (b) the predictions were not actually novel and risky, or (c) they were not deployed essentially (Alai 2014b).

Condition **(2')** is designed precisely to exclude cases of type (c), and I shall offer six examples in section 9.4 and three in section 9.5. Therefore, if **H** fulfills **(2')**, it must be *completely* true. But if so, we don't need *in addition* to exclude that there is any *incompatible* assumption **H'** from which **NP** could be essentially derived, because any **H'** incompatible with **H** would be (at least partly) false, hence it could not (except by a miracle) be essential in deriving **NP**: **H'** could play a role in deriving **NP** only if it had a completely true part **H'*** sufficient to derive **NP**, and obviously **H'*** would be compatible with **H**, since both would be completely true.

This is a nutshell account of why my condition **(2')** is enough to ensure the required uniqueness of **H**, but I shall explain it in more detail in section 9.6. This account, anyway, entails an empirical claim: in the history of science there have never been a couple of *incompatible* hypotheses both essential in deriving the same novel risky prediction. So, my proposal is empirically testable: it will be falsified if any such couple is found, but confirmed if it is not.

9.4. How the Essentiality Condition Rules out Lyons's Purported Counterexamples

Here is how my condition **(2')** rules out as inessential five false assumptions which according to Lyons were deployed to derive novel predictions (the first three are more extensively discussed in Alai 2014b, §7). Three further examples will be discussed in section 9.5. Yet another striking case is Arnold Sommerfeld's 1916 prediction of the fine structure energy levels of hydrogen: for decades it was thought to be derived from utterly false assumptions, and so considered "the ultimate historical challenge" to the NMA, for even physicists called Sommerfeld's success a "miracle." But Vickers (2018) shows that the wrong assumptions were not essential in deriving it, since the true part of Sommerfeld's premises was enough.

9.4.1 Dalton

Lyons argues that Dalton derived his true *Law of Multiple Proportions* (**LMP**) from his false *Principle of Simplicity*:

> **PS**: "Where two elements A and B form only one compound, its compound atom contains 1 atom of A and 1 of B. If a second compound exists, its atoms

will contain 2 atoms of A and 1 of B, and a third will be composed of 1 of A and 2 of B, etc." (Lyons 2002, 81; Hudson 1992, 81).

But this false **PS** has a *true* part: a weaker principle we might call "*of Multiple Quantities*":

PMQ: The quantity of atoms of B combining with a given number of atoms of A is always a multiple of a given number.

Besides, **PMQ** is enough to derive the *Law of Multiple Proportions*:

LMP: The weights of one element that combine with a fixed weight of the other are in a ratio of small whole numbers.

Therefore, although actually employed in deriving **LMP**, the false **PS** was not essential, since its true part **PMQ** was sufficient to derive **LMP**.

9.4.2 Mendeleev

Mendeleev derived his predictions of new elements from his false *Periodic Law*:

PL: atomic *weights* determine the chemical properties of elements.

PL is false, since chemical properties are rather determined by atomic *numbers* (Lyons 2002, 80, 84). However, **PL** entails that

Q: an atomic quantity *approximately proportional to atomic weight* determines chemical properties (where Mendeleev thought this quantity was atomic weight itself, while we know it is the atomic *number*).

Now, **Q** is a part of **PL**, true, and sufficient to derive Mendeleev' predictions. So, although actually employed, **PL** was not essential, while the essential assumption, **Q**, was true.

9.4.3 Caloric

Lyons (2002, 80) points out that the caloric theory truly predicted that

ER: The rate of expansion is the same for all gases.

But **ER** was derived from a number of false claims, viz.:

(1) heat is a weightless fluid called caloric;
(2) the greater the amount of caloric in a body, the greater its temperature;
(3) gases have a high degree of caloric;
(4) caloric, being a material itself, is composed of particles;
(5) caloric particles have repulsive properties which, when added to a substance, separate the particles of that substance;
(6) the elasticity of gases is caused by this repulsive property of caloric heat particles (Carrier 1991, 31).

However, these false claims were not *essential* in predicting **ER**. In fact, each of them includes a true part not referring to caloric, from which **ER** could still be derived, respectively

(1*) heat is weightless, and it expands like fluids;
(2*) the greater the amount of heat (whatever it consists in) in a body, the greater its temperature;
(3*) gases have a high amount of what heat consists in;
(4*) particles play a crucial role in the constitution of heat;
(5*) heat involves repulsive forces, which separate the particles of substances to which it is added;
(6*) the elasticity of gases is caused by the repulsive forces involved by heat.[8]

So, while the false (1)–(6) were used, they were not essential, since their parts (1*)–(6*) are sufficient to derive **ER** (and true, as far as we know).

9.4.4 Kepler

In (2006, 545–553) Lyons discusses a number of false premises from which Kepler derived (P1) that the Sun spins, (P2) that a planet's speed is highest at its perihelion and lowest at its aphelion, and (P3) his three laws. Further, since Kepler's laws were used by Newton, those false premises indirectly yielded also the successes of Newton's gravitation theory, including Neptune's discovery. An exhaustive analysis of Lyons's discussion would require a different paper. Here I can only sketch how my essentiality condition rescues the **NMA** even in this case.

Kepler's false assumptions were:

[8] See Psillos (1999, 115–130) for an extensive presentation of this strategy.

(#1) the planets tend to rest, and they move only when forced to move;
(#2) the Sun has an *Anima Motrix* which spins it around its axis;
(#3) this spinning is transmitted to the planets through the Sun's rays, which push the planets in their orbits (thus the Sun exerts on the planets a *directive* rotational force, not an *attractive* one);
(#4) the force of the Sun's rays is inversely proportional to the distance from the Sun, like the intensity of their light.

The model of the universe described by (#1)–(#4) is crazy in the light of our mechanics (#1), metaphysics (#2), and physics (#3), but it was not in Kepler's time. Moreover, it was a natural abductive conclusion from two facts known at that time:

(a) the planets move around the Sun approximately on the same plane and in the same wise;
(b) the order of their velocities is inverse to that of their distances from the Sun.[9]

These facts plausibly suggested a model analogous to mechanisms well known to Kepler, like a sling, where a central hub rotates, and through a strap transmits its motion to a peripheral body, which is then set into circular motion as well. In Kepler's model the hub is the Sun, the peripheral body is a planet, and the strap represents the Sun's rays. The sling has just one projectile, while the planets are many; but there are similar earthly mechanisms which preserve this analogy, like wool winders, wheels to lift water, or clock gears (although the sling model is better from the point of view of fact (b), for (initially) the projectile rotates more slowly than the hand, and the more slowly the longer the strap is).

The predictions (P1) that the Sun spins and (P2) that the velocity of the planets is highest at the perihelion and lowest at the aphelion, are immediate consequences of Kepler's model (#1)–(#4). Of course the model is wrong, for his abduction from the known facts (a) and (b) included some wrong guesses. But it is also unnecessarily strong as an explanation of those facts. Moreover, the wrong part was not *essential* to Kepler's novel predictions or to the discovery of his three laws, since the model had a weaker core which was both true and sufficient to derive them. Kepler might have restricted his assumptions to that weaker core if

[9] This was already clear to Copernicus, who knew the times of the orbits and the relative distances from the Sun, hence the relative lengths of the orbits. Kepler also knew that the times grow more than the lengths, hence the velocity diminishes as the distance from the Sun increases. This suggests that the planets be driven by a force emanating from the Sun, so varying inversely with the distance from it (Koyré 1961, II, §1). Kepler thought that the light varies inversely with the distance from its source (rather than with its square), hence he thought the Sun's motrix force did the same.

he had been more cautious in extrapolating from (a) and (b); but unfortunately there were no technological analogies, like the sling, the wool winder, etc., to suggest to him such a model. This true core is the following:

(#5) the solar system moves around the centre of the Sun as a coherent but non-rigid disk;
(#6) its coherence is *mainly* due to *a* force exerted by the Sun on the planets (Kepler thought it was a *directive* rotational force, and the only one in play; while we know it is the *attractive* gravitational force, supplemented by the planets' own attractive forces and the effects of inertia);
(#7) the velocities of the planets are also *proportional* to the same force (Kepler thought this happens because that force is responsible for their motion, by overcoming their tendency to rest; instead we know this happens because their motion is due to their tangential inertia, which must be equal to that (gravitational) force, otherwise the planet would either fly off on its tangent, or collapse on the Sun);
(#8) that force is in *an* inverse relation with the distance from the Sun (for Kepler that relation was $F=1/d$, we know it is $F=1/d^2$).

I am not saying that (#5)–(#8) by themselves constitute a viable model of the universe: we get one only by supplementing them either by the rest of Kepler's assumptions or by the rest of our assumptions[10] (in the first case, of course, we get a viable, but false model, in the second case a true model). For instance, we wouldn't get a viable model just by adding to (#6) that the force is an attractive one, without adding the planets' inertia, for then the planets would collapse on the Sun. This is probably why Kepler himself, after considering the idea of an attractive force, rejected it (Lyons 2006, 547). I am only claiming that (#5)–(#8) are a core shared by two otherwise very different models; that as far as we know they are true; and that they are enough to derive Kepler's novel predictions (P1)–(P3).

In fact, (P1) the spinning of the Sun is already implicit in (#5); (P2) that a planet's speed is highest at its perihelion and lowest at its aphelion is implied by (#7) and (#8); finally, (P3) Kepler's three laws are just kinematic laws, and once given (#5) they were found mainly by making hypotheses on which curves and functions would fit the data and checking them back with the data; but their discovery was helped by Kepler's revolutionary intuition that the orbits must not be found by geometrical models alone, but reasoning on the forces which govern them, as suggested by (#7) and (#8) (Hoskin 1999, ch. V). So, there were no miracles: the true model (#5)–(#8) is part of the false model (1#)–(#4) and

[10] Some of which are mentioned within parentheses after (#6), (#7), and (#8) respectively.

enough to generate (P1)–(P3); therefore the model (#1)–(#4) was not essential by condition **(2')**, hence the **NMA** does not commit to it.

For the sake of simplicity, in (#6), (#7), and (#8) I spoke of a *force*, as if Newton's gravitational theory were true; but, one could object, General Relativity shows that there are no gravitational *forces*, and the planets' motion is due to the curvature of spacetime caused by the Sun. The objection is fair, but it only shows that we must circumscribe more accurately the true core of Kepler's model, which we can do by generalizing (#6), (#7), and (#8) as follows:

(#5) the solar system moves around the centre of the Sun as a coherent but not rigid disk;
(#6*) its coherence is due to an effect (i.e., for us, the curvature of spacetime) which is mainly brought about by the Sun;
(#7*) the velocities of the planets are also proportional to the same effect;
(#8*) that effect is in an inverse relation with the distance from the Sun (hence so are the planetary velocities).

This restricted core is shared by Kepler, Newton, and Einstein and is still sufficient to derive Kepler's novel predictions. But what if someday General Relativity is replaced by a better theory and this by another one, etc.?[11] If we constantly weaken our hypotheses to keep them compatible with successive theories, won't we reduce them to mere descriptions of observed phenomena, and scientific realism to empiricism?[12] I don't think this pessimism is warranted, because history shows that each successive theory **T'**, while giving up part of the theoretical content of the earlier theory **T**, adds some new theoretical content. Even if in turn **T'** is superseded by **T''** and some of its content is given up, it is quite possible that (a) some of the original content of **T** is preserved even in **T''**, and (b) the theoretical content of **T'** preserved in **T''** is even larger than that of **T** preserved in **T'**.

For instance, in the case discussed here, the directive force has been replaced by the attractive force, and this by a curvature of spacetime, but all of them are unobservable entities or properties. Each theorist while discarding some theoretical assumptions adds some new ones (Newton adds inertia and gravitation force, Einstein spacetime curvature). Tomorrow it might no longer be spacetime curvature, but something else; but whatever it is, the Sun will have a major role in it, and it will have an inverse relation with the distance from the Sun and a direct relation with the velocities of the planets. Besides, each new model is probably closer to the truth, because it has a better fit to the data and greater predictive

[11] I deny that our current theories are completely true and will never be rejected: Alai (2017).
[12] I owe this objection to an anonymous reviewer.

power. Thus our theoretical knowledge (our theoretical true justified beliefs) increases over time. While acknowledging that our beliefs are not yet "the whole truth and nothing but the truth," we can trust to be on the right track. I shall further discuss similar worries in §§9.7.3–9.7.5.

9.4.5 Neptune

Finally, Lyons claims that the following false claims were used in the prediction of Neptune (2006, 554):

(1) the sun is a divine being and/or the center of the universe (Kepler);
(2) the natural state of the planets is rest;
(3) a non-attractive emanation coming from the sun pushes the planets in their paths;
(4) the planets have an inclination to rest, which resists the solar push, and contributes to their slowing speed when more distant from the sun;
(5) the planets are pushed by a "directive" magnetic force;
(6) there exists only a single planet and a sun in the universe (Newton);
(7) each body possesses an innate force, which, without impediment, propels it in a straight line infinitely;
(8) between any two bodies there exists an instantaneous action-at-a-distance attractive force;
(9) the planet just beyond Uranus has a mass of 35.7 earth masses (Leverrier)/50 earth masses (Adams);
(10) that planet has an eccentricity 0.10761 (Leverrier)/0.120615 (Adams);
(11) the longitude of that planet's perihelion is 284°, 45' (Leverrier)/299°, 11' (Adams), etc.

However, (1)–(5) were no longer held in the 18th century. They were involved in Neptune's discovery only to the extent that they were involved in the discovery of Kepler's laws, which in turn were used by Newton. But I argued that (1)–(5) were not involved *essentially* in Kepler's discoveries. (9)–(11) concern Neptune, so they could not be premises from which Neptune's existence or location was derived: they are only (partially wrong) parts of a global hypothesis on the planet's existence and behavior. (6) was practically true in this context, given the negligible influence of the other planets. To my knowledge (7) was never held by anybody; anyway, it entails the true claim that inertial motion continues indefinitely in a straight line. Finally, I just argued that (8), gravitational *force*, was not essential in Newton's predictions. So, even Neptune ceases to be a counterexample to **DR**.

9.5 Essentiality Cannot be Detected Prospectively

For deployment realists the components of discarded theories which were essentially involved in novel predictions are most probably true; but Stanford (2006), Votsis (2011), and Peters (2014) argue that it is not enough to identify these components just *retrospectively*, i.e., as those preserved in current theories. Rather, they must be identified *prospectively*, from the viewpoint of their contemporaries. In fact, deployment realists claim that the hypotheses of past theories still preserved today are true because they were essential. Thus they resist the Pessimistic Meta-Induction (**PM-I**), maintaining that both past and current theories are at least partly true. But explaining that those hypotheses were essential because they are preserved today, would be explaining that they are true because they are preserved today, so begging the question of the truth of current theories. In fact, since a hypothesis **H** is preserved today because we believe it is true, saying that it is true because it is essential, and it is essential because it is preserved today, would be saying it is true because we believe that it is true.

Equally, deployment realists claim that the now rejected hypotheses deployed in novel predictions were not essential, hence the **NMA** did not commit to their truth. In this way they block Laudan and Lyons's meta-*modus tollens* (**M-MT**). But explaining that they were not essential because they are not preserved today would be stipulating that the **NMA** warrants all and only the hypotheses accepted today. Thus the realist's defense against both the **PM-I** and the **M-MT** would be circular.

This is why it would seem desirable that the components of past theories which are essential for their novel predictions are identified prospectively; moreover, Votsis (2011), Peters (2014), Cordero (2017a, 2017b), and others claim that such prospective identification *is* possible. On the opposite side, Stanford (2006, 167–180; 2009, 385–387) claims this is impossible, therefore the arguments for **DR** are circular.[13]

I take an intermediate position: any time we believe a hypothesis is essential, it may in fact be, but neither at that time nor later can we ever be certain that it is. Yet, I argue that this does not make the arguments for **DR** circular. That we cannot ever be certain is shown by many hypotheses which were firmly believed by past scientists because they appeared to be essential to certain novel predictions, but subsequently turned out to be inessential, in fact false.

For instance, Fresnel and Maxwell derived various novel predictions from the hypothesis that

> **AV**: aether vibrates.

[13] See also Vickers (2013, 207).

Today we know that **AV** was inessential in these derivations, because it can be substituted by its weaker consequence

VM: there is a vibrating medium.[14]

Fresnel and Maxwell could easily have seen that **VM** was a proper part of **AV** and enough to derive those predictions. However, they didn't consider **VM** (probably they didn't even think of it), no doubt because they presupposed that

P_1: all mediums are material,

and/or that

P_2: all waves are produced by the oscillations of particles.

Hence, given their presuppositions, *any* vibrating medium was a *material* medium *composed of particles* (i.e., either water, or air, or *aether*). Therefore from their viewpoint, **AV** was required by their predictions as much as **VM**.

Again, Laplace predicted the speed of sound in air starting from the false hypothesis that

H: the propagation of sound is an adiabatic process, in which some quantity of caloric contained by air is released by compression.

Psillos (1999, 119–121) argues that **H** was not essential to Laplace's prediction, which could also be derived from a part of **H**, viz.

H*: the propagation of sound is an adiabatic process, in which some quantity of latent heat contained by air (whatever be the nature of heat) is released by compression.

However, at that time adiabatic heating could only be explained as the disengagement of caloric from ordinary matter, caused by mechanical compression (Chang 2003, 904), for it was presupposed that

P_1: gases can be heated without exchanges with the environment only if they contain heat in a latent form,

and

[14] Today we call it electromagnetic field.

> **P_2**: only material substances can be contained by material substances in a latent form.

But the material substance of heat was just caloric: therefore, given **P_1** and **P_2**, **H*** entailed **H**, hence **H** too had to be considered essential.

Generalizing, we might say that at any time in the history of science, when a novel prediction **NP** is derived from an assumption **H**, there may be an assumption **H*** which is proper part of **H** and sufficient—from a purely logical viewpoint—to derive **NP**; however, given certain current explicit or implicit presuppositions **PRS**, **H** may "appear to be conceptually or metaphysically entailed by" **H*** (Vickers 2016, §4), hence scientists may falsely believe that **H** is essential to **NP**. At a later time, however, the advancement of science may show that the **PRS** are false, so scientists can (retrospectively) recognize that **H** was inessential, after all. This recognition is facilitated if meanwhile experimental or theoretical doubts have arisen about **H**, so that scientists started wondering whether **H** was really necessary. Yet, this recognition is logically independent of those doubts, as it follows already from the falsity of the **PRS**; hence, although retrospective, it is not circular.

Vickers, however, suggests that although prospectively we cannot ever be certain that **H** is essential to a prediction, *in some cases* we could (still prospectively) recognize it is *not* essential, although "metaphysically or conceptually" required by other considerations. This may (i) suggest the realist to "restrict her commitments to what is *directly confirmed* by the predictive successes" (i.e., **H***), and (ii) supply "a worthwhile heuristic," i.e., show which commitments should be abandoned first in case of empirical or theoretical refutations (i.e., **H**).

For instance, Bohr predicted (**NP**) the spectral lines of ionized helium by assuming that

> **H**: The electron orbits the nucleus only on certain specific orbital trajectories, each characterized by a given quantized energy.

H turns out to be false, but **NP** could have been derived by the weaker hypothesis that

> **H***: The electron can only have certain, specific, quantized energy states.

Moreover, says Vickers, Bohr could have seen (prospectively) that **H** was inessential, because it entailed **H***, which was enough to derive **NP**. But he still held **H** because "it may have been inconceivable at the time to think that electrons could have quantized electron energies *without* having associated quantized orbital trajectories (cf. Stanford 2006, 171)". Nonetheless, Bohr might have distinguished between his reason for believing **H*** (i.e., the essentiality of **H*** in predicting **NP**), and his reason for believing the content of **H** exceeding **H*** (i.e., his presupposition that electrons could have quantized energies only by having orbital trajectories). Thus he might have realized that "the latter were not as secure as the former" (Vickers 2016, §4).

I agree, but with the following qualifications:

(α) in many cases prospective recognition of inessentiality will be impossible;
(β) essential hypotheses cannot be identified with certainty, neither prospectively (as argued by Stanford) nor retrospectively;
(γ) however, this is not a problem, because the **PM-I** can be blocked even without identifying essential hypotheses, and the **M-MT** can be blocked even by identifying inessential hypotheses *retrospectively*.
(δ) at any rate, recognizing inessentiality (especially prospectively) is not a task for (realist) philosophers, but for scientists;
(ε) even when inessentiality can be recognized prospectively, this does not help the realist defense against the **PM-I** and **M-MT**.

Here are my arguments:

(α) Perhaps in Bohr's case the extra-content of **H** could be distinguished from **H***; but in other cases it may consist in principles so obvious or deeply entrenched to pass unnoticed, such as the principles of conservation of energy and mass, isotropy and homogeneity of space, physical causal closure, etc. We usually presuppose so many principles of this kind, that even by paying close attention one cannot be certain to have ruled out all of them. Moreover, it may be extremely difficult to imagine an assumption **H*** that still entails **NP** once all of them have been discarded.

Besides, Vickers grants that **H** may be "conceptually" entailed by **H***, i.e., that the contemporaries may lack the conceptual resources needed to distinguish **H*** from **H**. For instance, those who used Newtonian *forces* to predict Neptune couldn't even conceive any "effect" (such as the curvature of spacetime) which could cause the planetary motions except a *force*. Therefore, when **H** *seems* to be essential in deriving **NP**, it may be impossible (either prospectively or retrospectively) to see that it is not.

(β) A fortiori, we cannot ever be certain that there is *no* weaker component **H*** from which **NP** could have been derived. Therefore, we cannot ever be certain that **H** is essential, neither prospectively (as Stanford claims) nor retrospectively. This is just natural, for if we could identify essential components we would know that they are *completely* true. But if so, we could anticipate scientific progress much more than we actually can: while trusting that current theories are largely true, scientists grant that some of their assumptions are probably wrong, but only future research will tell which ones.

Besides, the **PM-I** is probably right that no theory or component older than 100 years or so was *completely* true (but not that none was *at least partly* true). Cordero, Peters, and Votsis are also right that we know when a hypothesis **H** is "true," if this means "*at least partly* true": but even if deployed in novel predictions, **H** may have been inessential, and even if otherwise well supported it can be partly false, and sooner or later replaced by another more completely true.

The progress of research allows us to drop more and more false presuppositions. Thus we may discover, retrospectively, that **H** was inessential, for it had some proper part **H*** which (i) can be true even if **H** is false and (ii) is sufficient to derive **NP**. That weaker **H*** may also be actively sought for, if **H** encounters empirical or theoretical refutations, which suggest that it cannot possibly be essential (since essential components are completely true). In this case (still retrospectively), one can recognize that **H** was inessential even before identifying its weaker substitute **H***.

(γ) Neither the impossibility of recognizing essentiality, nor the difficulties in prospectively recognizing non-essentiality are a problem for deployment realists, because these recognitions are not necessary to resist the **PM-I** and the **M-MT**. In fact, *as soon as* a *risky novel* prediction **NP** derived from **H** is confirmed, we know (prospectively, and a fortiori retrospectively) that, short of miracles, there is *some* truth in **H**. This is enough to refute the **PM-I**'s claims that both past and present theories are completely false.

The **M-MT**, in turn, is always used *retrospectively*, therefore only a retrospective recognition of inessentiality is needed to resist it: if at time **t** hypothesis **H** was firmly believed because considered essential to derive **NP**, but at **t'** it is shown to be false, the **M-MT** uses it as a counterexample to the **NMA**. But realists can reply by two moves: (i) *arguing* that probably **H** was inessential, for only by a miracle could **NP** have been derived from a completely false assumption; (ii) *showing* **H** was inessential, by identifying a proper part **H*** sufficient to derive **NP**, and the false presuppositions **PRS** that at **t** prevented us from distinguishing **H** from **H***. Both moves are retrospective (at t' or at a later t"), yet sufficient to block the **M-MT**. Besides, move (ii) is independent of the refutation of **H**, therefore the realist's defense is not circular, as Stanford claims.

(δ) Even when it is possible, the recognition of non-essentiality (especially prospectively) is a task for scientists: philosophers just don't have the necessary expertise. The burden of scientific realists is arguing that certain criteria (e.g., essential deployment) can justify our beliefs in the at least partial truth of theories

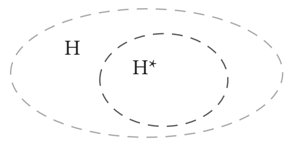

Figure 9.1

or hypotheses. But Fahrbach (2017) asks too much when he requires that they also apply those criteria to actual research, so teaching practitioners which are the working hypotheses and which are the idle parts in their theories, urging changes, suggesting directions of research, etc.

Therefore, *qua* philosophers, realists need not be committed to any particular theory or hypothesis, not even to the best current ones: as argued by Smart (1963, 36), this is a task for scientists. At most, they may argue that science is convergent, hence, *in general*, current theories are *probably* more largely true than past ones. Therefore, *pace* Stanford (2017), they need not be more conservative than anti-realists. Actually, since they set for hypotheses a higher standard than anti-realists—truth, rather than empirical adequacy or the like—for them it is even more likely that any particular hypothesis fails to reach that standard, hence must be substituted by a better one.

(ε) The discussion at (γ) shows that even when inessentiality can be recognized prospectively, that is superfluous in resisting the **PM-I** and the **M-MT**. In fact, the at least partial truth of **H** can be recognized prospectively, even if **H** is not essential, and this is enough to block the **PM-I**. Besides, prospective recognition of inessentiality is superfluous against the **M-MT**, because the latter is always used retrospectively.

9.6 A More Complete Account of the Dispensability of Psillos's Condition (2)

Here I explain in more detail why my condition **(2')** is enough and we can do without the greater complexity of Psillos's condition **(2)** or its possible strengthening implicitly suggested by Lyons. When Psillos excludes any *other* hypothesis **H***', **H*** may be compatible or incompatible with **H**, and in different mereological relations to it. The following cases are included:

(**a**) **H*** is proper part of **H**. This is precisely the kind of alternative hypothesis **H*** excluded by my condition **(2')** (Figure 9.1)
(**b**) **H** is a proper part of **H*** (Figure 9.2)

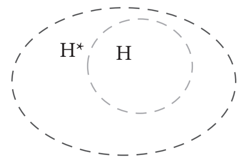

Figure 9.2

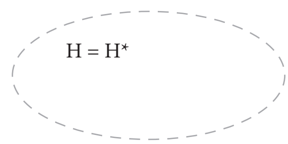

Figure 9.3

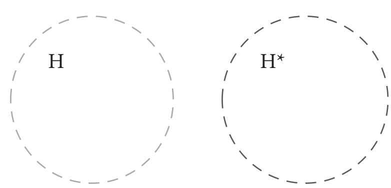

Figure 9.4

(**c**) **H*** and **H** coincide (Figure 9.3)
(**d**) Neither hypothesis is part of the other, and they have no common content (Figure 9.4)
(**e**) Neither hypothesis is part of the other, since they have some common content but two alternative extensions (Figure 9.5)

Now, Psillos's condition (**2**) excludes alternative hypotheses **H*** of all these five kinds, and here the troubles arise, while my condition (**2'**) excludes only alternative hypotheses **H*** of kind (**a**), and I claim that this is enough for deployment realists. In cases (**b**), (**d**), and (**e**) **H*** has some content outside **H**; so, in these cases **H*** might be incompatible with **OHs** and **AA**, and Lyons implicitly suggests they should be excluded along with the cases in which **H*** is compatible with **OHs** and **AA**. But we don't need to worry about any of these cases, and here is why:

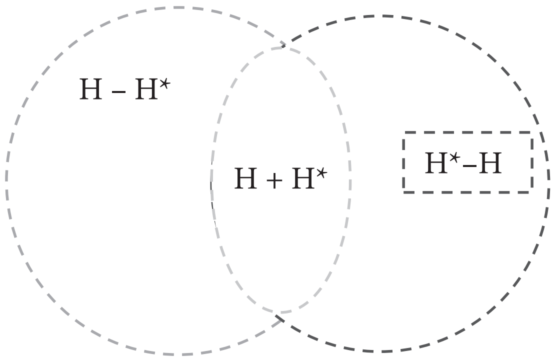

Figure 9.5

In case **(b) H** is weaker than and part of **H***. Therefore the existence of **H*** does not make **H** inessential, so there is no need to exclude it. In fact, there have always been such unnecessarily rich alternatives to the true hypothesis (like e.g. the ether theory with respect to the simple assumption of Maxwell's equations, or the aforementioned demi-gods theory with respect to Newton's actual theory). From a purely logical viewpoint, for any **H** there are infinite such **H***, but they do not make **H** inessential.

In case **(c) H*** coincides with **H**, so again it does not make **H** inessential and we don't need to exclude it.

In case **(d) H** and **H*** have no common content, and this may happen in two ways:

(d1) **H** and **H*** are compatible, hence, they may be both true. In general it is quite possible that a prediction is reached by independent inferential routes starting from different hypotheses. This is just natural if we conceive the world as a tightly connected whole, in which each phenomenon, at each scale, is causally linked to many other phenomena even at different scales. For instance, Perrin famously derived the same value for Avogadro's number reasoning from tenets of chemistry, thermodynamics, electrical theory, the theory of Brownian motions, and others. In such cases, each hypothesis $\mathbf{H}_i$ deployed in a different inferential route to **NP** is not "essential" in the sense that without $\mathbf{H}_i$ one could not have predicted **NP** (for **NP** *could* have been predicted through some different inferential routes involving some alternative hypothesis $\mathbf{H}_j$). However, $\mathbf{H}_i$ can be *essential* in the sense that **NP** could not have been predicted along that particular inferential route without $\mathbf{H}_i$, i.e., in the sense that $\mathbf{H}_i$ cannot be substituted in the same inferential route by one of its proper parts. This is precisely what my condition **(2')** requires, and it is all we need from a criterion of essentiality, for it is enough to warrant the complete truth of $\mathbf{H}_i$ (for we saw that, short of miracles, a hypothesis deployed in predicting **NP** cannot be completely false, and it can be partly false only if its false part is dispensable in the derivation). Therefore we need not exclude case **(d1)**, for in this case **H** is essential in the required sense if it fulfills **(2')**. By the way, since **H** entails **NP** only in conjunction with **AA** and **OHs**, also **AA** and **OHs** are essential if they satisfy **(2')**.[15]

(d2) Otherwise, **H** and **H*** may have no common part and be incompatible. One might think that we need to exclude this case: for if we had two incompatible hypotheses both predicting **NP**, we couldn't know which of them is true.[16] However, as explained in my nutshell account at section 9.3, this case is already excluded by the *risky* character of **NP**. In fact, as argued earlier, it is practically impossible to find a *completely* false hypothesis entailing a risky novel prediction. But if **H** is essential in the sense of **(2')**, it must be completely true. Therefore, if **H*** is incompatible with it, we know that (i) **H*** is at least partly false and (ii) it cannot possibly predict **NP**, for it contradicts an assumption (i.e., **H**) which is essential to that prediction.[17] Therefore, my condition **(2')** (together with the risky character of **NP**) already excludes that any hypothesis incompatible with **H** also predicts **NP**.

Finally, **case (e): H** and **H*** partially overlap. Like in case **(d)**, here too we must distinguish:

(e1) If **H** and **H*** are compatible, it is possible, though unlikely, that they are both are essential, hence certainly true, as explained in case **(d1)**. Otherwise, the essential assumption is neither **H** nor **H***, but some part of their common content.

(e2) If **H** and **H*** are incompatible (like e.g. Rutherford's and Bohr's theories of the atom), neither of their non-overlapping parts is required by the derivation: in fact, **NP** is predicted by **H** without using **H*-H**, and it is predicted by **H*** without using **H-H***. Therefore both

[15] Then we know they are true (unless **NP** was used in building them).

[16] Besides, we would have a counterexample to the **NMA**, for there would be at least one false hypothesis allowing us to derive a novel prediction.

[17] Instead, if **H** is not essential in my sense, it consists of an essential part $\mathbf{H}_e$ and an inessential part $\mathbf{H}_i$. If **H*** also has a role in predicting **NP**, it may consist in an essential part $\mathbf{H^*}_e$ and an inessential part $\mathbf{H^*}_i$. Of course $\mathbf{H^*}_e$ cannot contradict $\mathbf{H}_e$, but $\mathbf{H^*}_i$ may contradict $\mathbf{H}_i$, in which case **H** and **H*** are incompatible, yet have both a (non-essential) role in predicting **NP**, and are both partly true. Instead $\mathbf{H}_e$ and $\mathbf{H^*}_e$ are two essential but compatible hypotheses as considered in case **(d1)**.

> **H** and **H*** have a superfluous part, hence neither is essential. If **H** were essential in my sense, it would be practically impossible that **H*** predicted **NP**, for it would contradict an assumption which is needed to predict it. So, even in case **(e2)** condition **(2')** is enough to warrant the truth of **H**.

Summing up, my condition **(2')** is enough to guarantee that a hypothesis **H** is *essential* in the derivation of a novel prediction **NP**, in a sense which warrants its complete truth, and we don't need Psillos's more complex and troublesome condition **(2)** or its strengthening implicitly suggested by Lyons.

9.7 Objections and Answers

Here I reply to some potential objections.

9.7.1 The "False May Entail True" Objection

Since a falsity may entail a truth, even false theories might make novel predictions.

Answer:

As argued at the beginning of section 9.1, in the Hyperuranion of possible theories there are some with wholly false theoretical assumptions which make true novel and improbable predictions (**F**$^{\mathbf{NIP}}$-theories). But their rate is exceedingly small, hence it is extremely unlikely that scientists pick one of them (because scientists look for true theories, hence they could find **F**$^{\mathbf{NIP}}$-theories only by chance). On the opposite, *all* true theories have true consequences, and those sufficiently fecund have true *novel* consequences. Granted, true (and fecund) theories are much fewer than false ones; however, they are not found *by chance*, but *on purpose* and through *reliable methods* (White 2003; Alai 2016, 552–554; 2018, §III).

9.7.2 The Disjunctive Objection

Your condition **(2')** cannot be fulfilled, hence no hypothesis is essential by your criterion, because any hypothesis **H** always entails some weaker claim **H'**, like e.g.,

Disjct: H or the Moon is made of cheese.

Answer:

First, condition **(2')** excludes that **NP** is entailed by any hypothesis which is *part* of **H** (in Yablo's sense), not just by any hypothesis entailed by **H**. In fact, **Disjct** is not *part* of **H**. Second, disjunctions are weaker than their disjuncts. Therefore, if **H** entails **NP**, in most cases **H***: «**H** or **H'**» does not. Hence, in such cases **H** fulfills **(2')**. For instance, Newton's Gravitation Law entailed the existence of Neptune; but the disjunction of this law with 'The Moon is made of cheese' does not. Therefore in certain cases **NP** cannot be entailed by any hypothesis which is part of **H**, hence **H** is essential.

In particular instances the disjunction of **H** with another hypothesis **H'** might still be strong enough to entail **NP** (Vickers 2016, §3). Obviously, however, this can happen only if even **H'** alone entails **NP**. But in this case **H** and **H'** must be in one of the five mereological relations examined in section 9.6, and I argued that in each of them condition **(2')** is enough to decide whether **H** (and **H'**) is essential or not.

9.7.3 The "Theoretician's Dilemma" Objection 1

The author's condition (**2'**) seems to make realism redundant for purely logical reasons. A statement is derivable from a set of statements iff the content of that statement is already included in that set of statements. The only proper part of that set of statements that's required to derive the statement is in fact the statement itself. Thus, a novel prediction **NP** needs only **NP** to be entailed. But surely we want to be realists about more things than the content corresponding to **NP**.[18]

Answer:

Most frequently and typically **NP** is not part of **H**, because, as explained by Vickers[19] and others, it is not entailed by **H** alone, but only in conjunction with certain other hypotheses **OHs** of **T**, and various auxiliary assumptions **AA**. Therefore (**2'**) does not allow us to dispense with **H** in favor of **NP**.

But the objection might become that since **H&OHs&AA** entails **NP**, we can dispense with **H&OHs&AA** in favor of **NP**. This is reminiscent of Hempel's (1958) "theoretician's dilemma": theoretical hypotheses **H** are required to connect the observable initial conditions **IC** to the observable final conditions **FC**, and they are justified only if they succeed in this, i.e., if **IC→H→FC**. But then why not drop **H** and just keep the empirical laws **IC→FC**? Hempel answered that this would work from a purely logical point of view, but not from an epistemological point of view: **H** does not predict only **IC→FC**, but many other phenomena as well, most of which we would never have discovered without assuming **H**.

[18] Objection raised by an anonymous referee.
[19] Vickers (2013, p. 202 footnote 9; 2016, footnote 8).

The answer here is similar: we would never have found **NP**, and many other predictions, unless **H&OHs&AA** were true. In fact, **NP** is novel and risky, hence it cannot have been found just by chance. Besides, the auxiliary assumptions **AA** are the consequences of many independent theories **T'**, **T''**. . . $\mathbf{T^n}$, typically dealing with matters different from **T**. Therefore, it would be a miraculous coincidence if even one of them were completely false, yet the conjunction of their consequences entailed **NP**. The only plausible explanation is that **T**, **T'**, **T''**. . . $\mathbf{T^n}$ are all at least partially true (Alai 2014c, §4). Besides, each of **T**, **T'**, **T''**. . . $\mathbf{T^n}$, in conjunction with still different assumptions, issues many other predictions, which are only explainable in the same way.

One might reply that guessing **n** true theories is even less likely than guessing **n** false theories with one true joint consequence. This is right, but guessing **n** false theories with one true, novel, and risky joint consequence is still too improbable to be a minimally plausible explanation of our frequent predictive successes. So, how was it possible to derive **NP** and so many other novel predictions? The only plausible explanation is that theories are not *guessed*, but found by a method which is reliable in tracking truth. Scientific realists have a reasonable account of why the scientific method is reliable, while anti-realists have never been successful in providing an account of how predictive success might be achieved by pure chance or "miracles" (Alai 2014a). Notice, here the truth-conduciveness of the scientific method is not a *petitio principii*, but the conclusion of my inference to the only plausible explanation in this paragraph (the second in this section).[20]

9.7.4 The "Theoretician's Dilemma" Objection 2

By your criterion no hypothesis will ever be essential, because whenever **H** together with **OHs** and **AA** entails prediction **P**, there is always a weaker hypothesis doing the same, viz.

H*: If **OHs&AA**, then **P**.[21]

Answer:

Again, here **H*** is not part of **H**, hence its existence does not make **H** inessential. For instance, suppose the following:

H: The atoll is loaded with deadly radioactive material;

[20] White (2003); Alai (2016, 552–554; 2018 §III).

[21] Objection raised in discussion by John Worrall, who cited a *topos* from the debates on confirmation.

OHs&AA: Carl lands on the atoll;
P: Carl will die;
H*: If Carl lands on the atoll he will die.

Of course **H*** together with **OHs&AA** entails **P**. However **H*** is not part of **H**, because there are truthmakers of **H*** (e.g., «There will be a disastrous *tsunami* on the atoll») not implied by any truthmaker of **H**. One relevant consequence is that while **H** explains **P**, **H*** does not. Even more importantly, in cases like this **H*** is a purely "empirical" conditional, connecting the initial conditions with an empirical consequence. Therefore, unless it were inferred from the "theoretical" hypothesis **H**, **H*** could be discovered only a posteriori, by witnessing **P**. Therefore **P** would not be novel, and **H*** would not have been deployed in a *novel* prediction.

9.7.5 The "Ramsey Sentence" Objection

Given that whatever follows from **H** also follows its Ramsey sentence ($\mathbf{RS^H}$), and that **H** implies $\mathbf{RS^H}$, wouldn't it follow that every hypothesis is dispensable in favor of its Ramsey sentence?[22]

Answer:

No: in general, the essential part **H*** of **H** does not coincide with $\mathbf{RS^H}$, because it contains both something more and something less than $\mathbf{RS^H}$. It must contain something more because $\mathbf{RS^H}$ would not allow to derive **NP**. For instance, assuming in Newton's gravitation law (**N**) the terms "force" and "mass" are theoretical and "distance" is observational, its Ramsey sentence is

$$\mathbf{RS^N}: \exists x, \exists y\left(\mathrm{x=Gy_1y_2/D^2}\right).$$

But $\mathbf{RS^N}$ is not enough to predict Neptune. We also need to say what kind of properties or relations are x and y, how they behave, etc.: we need an expanded Ramsey sentence like

$$\mathbf{RS^{NEx}}: \exists \mathrm{x}, \exists \mathrm{y}\left(\ldots.\mathrm{x}\ldots.\ \&\ \ldots.\mathrm{y}\ldots.\ \&\ \mathrm{x=Gy_1y_2/D^2}\right),$$

where '....x...' and '...y...' are various laws concerning x and y (which, in practice, characterize them respectively as force and mass).

[22] Objection raised by an anonymous referee.

But typically the essential part **H***contains also something less than $\mathbf{RS}^{\mathbf{H}}$ or $\mathbf{RS}^{\mathbf{HEx}}$. For instance, Newton's Gravitation Law is partly false (since there exists no gravitation force). Consequently, also $\mathbf{RS}^{\mathbf{N}}$ and $\mathbf{RS}^{\mathbf{NEx}}$ are partly false, hence, inessential. Therefore, in general, the essential **H*** is different from $\mathbf{RS}^{\mathbf{H}}$. It might still be representable by an expanded Ramsey sentence $\mathbf{RS}^{\mathbf{H^*Ex}}$ (a proper part of $\mathbf{RS}^{\mathbf{HEx}}$), but this wouldn't be a problem for realism, because:

(i) Like any Ramsey sentence, $\mathbf{RS}^{\mathbf{H^*Ex}}$ is not a merely empirical claim, it states the existence of unobservable entities, even without naming them.
(ii) In general, $\mathbf{RS}^{\mathbf{HEx}}$ will have a very complex structure, hence it might be practically impossible to write it out completely, so to *dispense* with **H**. Besides, as argued in section 9.5, identifying the essential part of hypotheses is very difficult, so even if we could spell out $\mathbf{RS}^{\mathbf{HEx}}$, often we would be unable to dispense with it in favor of its essential part $\mathbf{RS}^{\mathbf{H^*Ex}}$.
(iii) At any rate, my claim is that the **NMA** commits only to essential hypotheses, not that only essential hypotheses can be true. So, if the essential hypothesis turns out to be something like $\mathbf{RS}^{\mathbf{H^*E}}$, since Ramsey sentences characterize only partially the entities to which they refer, realists may hold that something stronger than $\mathbf{RS}^{\mathbf{H^*E}}$ is true.

9.7.6 The "Modus Ponens" Objection

By your criterion no hypothesis will ever be essential, because whenever **H** together with **OHs** and **AA** entails NP, numberless but intuitively irrelevant hypotheses do the same, for instance,

H**: (If God exists, then **H**) & God exists.[23]

Answer:

As explained above, **H**** is neither part of **H** nor implied by it. Therefore **H**** does not make **H** inessential.

Acknowledgments

I am grateful to the organizers of the "Quo Vadis Selective Scientific Realism?" conference in Durham (2017) for generously helping my participation with

[23] Objection raised in discussion by John Worrall, referring to another *topos* from the debates on confirmation.

a travel bursary. I received helpful comments and suggestions from Gustavo Cevolani, José Díez, Vincenzo Fano, Ludwig Fahrbach, Greg Frost-Arnold, Pablo Lorenzano, Giovanni Macchia, Flavia Marcacci, Matteo Plebani, Andrea Sereni, John Worrall, and especially Peter Vickers and Timothy Lyons. Thanks to Maria Grazia Severi and Peter Vickers for checking my English grammar.

This research was supported by the University of Urbino Carlo Bo through the *Contributo d'Ateneo progetti PRIN2015* and the DiSPeA Research Projects 2017 and by the Italian Ministry of Education, University and Research through the PRIN 2017 project "The Manifest Image and the Scientific Image" prot. 2017ZNWW7F_004. For this work, I profited from a leave of absence granted by the University of Urbino (Department of Pure and Applied Sciences).

References

Alai, M. (2012) "Levin and Ghins on the 'No Miracle' Argument and Naturalism," *European Journal for Philosophy of Science* vol. 2, n. 1, 85–110. http://rdcu.be/mTcs.

Alai, M. (2014a) "Why Antirealists Can't Explain Success," in F. Bacchini, S. Caputo and M. Dell'Utri (eds.) *Metaphysics and Ontology Without Myths*, Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 48–66.

Alai, M. (2014b) "Defending Deployment Realism against Alleged Counterexamples," in G. Bonino, G. Jesson, J. Cumpa (eds.) *Defending Realism. Ontological and Epistemological Investigations*, Boston-Berlin-Munich: De Gruyter, pp. 265–290.

Alai, M. (2014c) "Novel Predictions and the No Miracle Argument," *Erkenntnis* vol. 79, n. 2, 297–326. http://rdcu.be/mSra.

Alai, M. (2014d) "Explanatory Realism," in E. Agazzi (ed.) *Science, Metaphysics, Religion: Proceedings of the Conference of the International Academy of Philosophy of Science, Siroki Brijeg 24–24 July 2013* Milano, Franco Angeli, pp. 99–116.

Alai, M. (2016) "The No Miracle Argument and Strong Predictivism vs. Barnes," in L. Magnani and C. Casadio (eds.) *Model-Based Reasoning in Science and Technology. Logical, Epistemological, and Cognitive Issues*, Cham: Springer, pp. 541–556.

Alai, M. (2017) "Resisting the Historical Objections to Realism: Is Doppelt's a Viable Solution?" *Synthese* vol. 194, n. 9, 3267–3290. http://rdcu.be/mSrG.

Alai, M. (2018) "How Deployment Realism Withstands Doppelt's Criticisms," *Spontaneous Generations* vol. 9, n. 1, 122–135.

Carrier, M. (1991) "What is Wrong with the Miracle Argument?" *Studies in History and Philosophy of Science* vol. 22, 23–36.

Chang, H. (2003) "Preservative Realism and Its Discontents: Revisiting Caloric," *Philosophy of Science* vol. 70, 902–912.

Clarke, Steve, and Timothy D. Lyons (eds.) (2002) *Recent Themes in the Philosophy of Science. Scientific Realism and Commonsense*. Dordrecht: Kluwer.

Cordero, A. (2017a). *Retention, truth-content and selective realism*, in E. Agazzi (ed.) *Varieties of Scientific Realism*, Cham: Springer, pp. 245–256.

Cordero, A. (2017b) "Making Content Selective Realism the Best Realist Game in Town," *Quo Vadis Selective Scientific Realism?* Conference, Durham, UK 5–7 August 2017.

Fahrbach, L. (2017) "How Philosophy Could Save Science," *Quo Vadis Selective Scientific Realism?* Conference, Durham, UK 5–7 August 2017.

Hempel, Carl G. (1958) "The Theoretician's Dilemma: A Study in the Logic of Theory Construction," *Minnesota Studies in the Philosophy of Science* vol. 2, 173–226.

Hoskin, M. (ed.) (1999) *The Cambridge Concise History of Astronomy*, Cambridge: University Press.

Hudson, J. (1992) *The History of Chemistry.* London: Macmillan Press.

Kitcher, P. (1993) *The Advancement of Science*. New York: Oxford University Press.

Koyré, A. (1961) *La révolution astronomique*. Paris: Hermann.

Laudan, L. (1981) "A Confutation of Convergent Realism," *Philosophy of Science* vol. 48, 19–49.

Lipton, P. (1991) *Inference to the Best Explanation*, London: Routledge.

Lipton, P. (1993) "VI—Is the Best Good Enough?," *Proceedings of the Aristotelian Society* vol. 93, 89–104.

Lipton, P. (1994) "Truth, Existence, and the Best Explanation," in A. A. Derksen (ed.), *The Scientific Realism of Rom Harré*. Tilburg: Tilburg University Press, pp. 89–11.

Lyons, T. D. (2002) "The Pessimistic Meta-Modus Tollens," in S. Clarke and T. D. Lyons (eds.) *Recent Themes in the Philosophy Science. Australasian Studies in History and Philosophy of Science*, vol 17, Dordrecht: Springer, pp. 63–90. doi:10.1007/978-94-017-2862-1_4.

Lyons, T. D. (2003) "Explaining the Success of a Scientific Theory," *Philosophy of Science* vol. 70, 891–901.

Lyons, T.D. (2006) "Scientific Realism and the Stratagema de Divide et Impera," *The British Journal for the Philosophy of Science*, vol. 57, 537–560.

Lyons, T.D. (2009) "Criteria for Attributing Predictive Responsibility in the Scientific Realism debate: Deployment, Essentiality, Belief, Retention," *Human Affairs* vol. 19, 138–152.

Lyons, T.D. (2016) "Structural Realism versus Deployment Realism: A Comparative Evaluation," *Studies in History and Philosophy of Science* Part A vol. 59, 95–105.

Lyons, T.D. (2017) "Epistemic Selectivity, Historical Threats, and the Non-Epistemic Tenets of Scientific Realism," *Synthese* vol. 194 n. 9, 3203–3219.

Musgrave, A. (1985) "Realism versus Constructive Empiricism," in P. Churchland and C. Hooker (eds.) *Images of Science*, Chicago, Chicago University Press, pp. 197–221.

Musgrave, A. (1988) "The Ultimate Argument for Scientific Realism," in Robert Nola (ed.) *Relativism and Realism in Science*, Dordrecht: Springer, pp. 229–252.

Peters, D. (2014) "What Elements of Successful Scientific Theories are the Correct Targets for "Selective" Scientific Realism?" *Philosophy of Science* vol. 81, 377–397.

Plebani, M. (2017) "Aboutness for Impatients," *Academia.edu*, https://www.academia.edu/34295884/Aboutness_for_the_impatient.

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

Sankey, H. (2001) "Scientific Realism: An Elaboration And A Defense," *Theoria,* vol. 98, 35–54.

Smart, J.J.C. (1963) *Philosophy and Scientific Realism*. London: Routledge.

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

Stanford, K. P. (2009) "Author's Response," in "Grasping at Realist Straws," a Review Symposium of Stanford (2006), *Metascience* vol. 18, 379–390.

Stanford, P. K. (2017) "A Difference That Makes a Difference: Stein on Realism, Instrumentalism, and Intellectually Nourishing Snacks," *Quo Vadis Selective Scientific Realism?* Conference, Durham, UK 5–7 August 2017.

Vickers, P. (2013) "A Confrontation of Convergent Realism," *Philosophy of Science* vol. 80, n. 2, 189–211.

Vickers, P. (2016) "Understanding the Selective Realist Defence against the PMI," *Synthese* vol. 194, 3221–3232. https://doi:10.1007/s11229-016-1082-4.

Vickers, P. (2018) "Disarming the Ultimate Historical Challenge to Scientific Realism," *The British Journal for the Philosophy of Science* axy035, https://doi.org/10.1093/bjps/axy035

Votsis, I. (2011) "The Prospective Stance in Realism," *Philosophy of Science* vol. 78, 1223–1234.

White, R. (2003) "The Epistemic Advantage of Prediction over Accommodation," *Mind* vol. 112 n. 448, 653–683.

Wright, J. (2002) "Some Surprising Phenomena and Some Unsatisfactory Explanations of Them," in Clarke and Lyons (eds.), *Recent Themes in the Philosophy of Science.* Australasian Studies in History and Philosophy of Science, vol. 17. Dordrecht: Springer, pp. 139–153. https://doi:10.1007/978-94-017-2862-1_4.

Yablo, S. (2014) *Aboutness*. Princeton: Princeton University Press.