Studies in History and Philosophy of Science 104 (2024) 78–87

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Convergence strategies for theory assessment

Elena Castellani

Department of Humanities and Philosophy, University of Florence, via della Pergola 60, 50121, Firenze, Italy

ARTICLE INFO

Keywords:
Empirical and non-empirical theory assessment
Convergence arguments
Robustness
Consilience
Perrin
Extra dimensions in early string theory

ABSTRACT

This paper addresses the issue of the import of convergence arguments in theory assessment. A first part is devoted to making the point of the different types of strategies based on convergence, providing new distinctions with respect to the existing literature. Specific attention is devoted to robustness vs consilience arguments and one representative example for each category is then discussed in some detail. These are: (a) Perrin's famous robustness argument on behalf of the atomic hypothesis on the grounds of the concordance of thirteen different procedures to the same result for the Avogadro number; (b) the consilience argument motivating the trust in the viability of the extra-dimension conjecture in the context of early string theory. These two cases are expressly chosen in order to highlight possible differences, also including whether the convergence obtains in terms of empirical or theoretical procedures. Notwithstanding these various differences, in both cases the evaluation of the assessment strategy similarly depends, in a significant way, on how the convergence argument is interpreted, as shown in the final part of the paper.

1. Introduction

In current debates on the status of fundamental physics, a typical criticism to research fields at the frontiers of physics such as string theory or cosmic inflation is that they have more to do with pure mathematics (e.g., Hossenfelder, 2018), or even "fashion, faith and fantasy" (Penrose, 2016), than with traditional scientific methodology. The main reason for such criticism is the absence of empirical confirmation: the scenarios proposed in those theoretical frameworks are apparently so far away from the possibility of empirical testing that theory assessment grounded on empirical data does not seem a viable option. Thus, at least for the time being, one has to rely on assessment criteria that are not based on empirical testing: such as, for example, the "theoretical virtues" famously discussed by Kuhn (1970) for theory choice, or, more recently, the meta-empirical arguments individuated by Dawid (2013, 2021) for boosting the trust in the viability of a theory in the absence of empirical data. Such criteria, however, do not seem to be appropriate methodological tools for pursuing a line of research in investigating Nature, critics claim. Moreover, according to some of them, persevering under such motivations means "going astray", that is, abandoning "the" scientific method.[1]

In fact, the issue of the legitimacy of the criteria employed for theory assessment in scientific practice is more nuanced than some of the contenders in the above debate seem to assume. As has been variously noted in recent literature,[2] the debate (in its more mediatic form) has significantly suffered from not paying due attention to the subtleties of scientific methodology, as well as to the actual historical developments of the theories considered. Indeed, when examined under a more carefully detailed perspective, the effective deployment of theory building in fundamental physics shows a different story from what commonly depicted in the critical literature.[3]

A clear example is provided by considering those epistemic strategies which are common to both theoretical and empirical scientific practice. Precisely because they are shared, they provide an interesting perspective for discussing scientific methodology, especially when the debate is focused on contrasting empirical with non empirical cases. Particularly representative, in this respect, are the strategies for theory assessment which are based on *convergence* criteria. As we will see, notwithstanding the diversity of the convergence arguments one can envisage, the evaluation of the corresponding assessment strategy does not barely depends on such differences as, in particular, whether the convergence obtains in terms of empirical or theoretical procedures. The evaluation of the assessment process in these cases is a subtle task,

E-mail address: elena.castellani@unifi.it.

1 "How Beauty Leads Physics Astray" is the subtitle of Hossenfelder (2018), for example.

2 See, for example, the discussions in Dardashti et al. (2019).

3 On this point, van Dongen (2021) provides a historical informed analysis.

https://doi.org/10.1016/j.shpsa.2023.12.010
Received 4 July 2022; Received in revised form 14 December 2023; Accepted 14 December 2023

Available online 13 March 2024

and significantly depends on how the convergence argument is being interpreted, case by case.

Now, turning to the actual subject matter of this paper, what is intended by a convergence argument? In fact, there is a rich variety of convergence reasoning acting at the level of scientific practice, informing both theory-building processes and arguments for theory assessment. More precisely, under the big umbrella of "convergence methodology" different strategies can be included, involving such notions as *robustness*, *consilience*, *coherence* and *unification*. Given the various uses of these notions in today's literature on theory assessment, section 2 is devoted to survey the main arguments based on the convergence (or "concordance", as some prefer to call it) of significant features or results. In particular, a new, historically oriented analysis of robustness and consilience arguments is provided. Section 3 discusses in some detail one representative example each for robustness and consilience. These are: (a) Perrin's famous argument on behalf of the atomic hypothesis on the grounds of the convergence of thirteen different procedures to the same result for the Avogadro number; (b) the convergence argument motivating the trust in the viability of the extra-dimension conjecture in the context of early string theory. The two cases are expressly chosen in order to highlight possible differences, also including whether the convergence obtains in terms of (theory-mediated) empirical procedures or of purely theoretical procedures (however physically motivated), besides the distinction between the types of the convergence arguments implied. In both cases, however, the trust in the theory or hypothesis involved is undoubtedly boosted on the basis of the convergence. In which way, exactly? The last section is devoted to address this point by examining the kind of epistemic strategy at work in each case, and the related interpretative issue.

2. Varieties of convergence

Convergence-based reasoning is widespread in scientific practice, both in empirical and theoretical cases. Various types of convergence can be singled out, depending on the context, the intended aim and the convergent feature one is dealing with. Broadly speaking, convergence in science indicates that some relevant elements in scientific activity - experimental or theoretical results, methods, models or even theories – turn out to be the same thing or to be strongly related to each other. On the grounds of the existence of such a convergence, conditions for a successful convergence-based argument are commonly held to be a) the existence of genuinely different starting points, and b) the variety and independence of the paths by means of which the convergence is obtained.[4] How a convergence is obtained characterises the kind of reasoning which can be based on it. Commonly, convergence resulting from varied evidence is used to build arguments for boosting trust in the theoretical developments involved. This is the kind of convergence-based arguments – shortly, *convergence arguments* (*CAs*) – the paper focuses on.

CAs cases are variously described and interpreted in the literature. An exemplary illustration of such a variety is provided by the philosophical discussion on the famous argument attributed to Jean Perrin for assessing the atomic hypothesis. The argument, apparently based on the convergence to the same empirical result of the thirteen procedures followed for obtaining Avogadro's number, is undoubtedly the most debated case of *CA*.[5]

While Cartwright (1983, pp. 84–86) interprets it as an inference to the most probable cause, Salmon (1984, p. 220) considers it as a type of common cause argument. Most often, especially in recent literature, it is discussed as a paradigmatic instance of a *robustness* argument, with further specifications depending on how robustness analysis is intended.[6] All the mentioned descriptions correspond to the same "orthodox robustness interpretation" according to Hudson (2020a, p. 196), who proposes instead a different understanding of the argument in terms of calibration (2020a) or analogical reasoning (2020b). Finally, Coko (2020a) provides a detailed discussion of Perrin's argument as a case of multiple determination.

Another representative example from the history of physics is Newton's theory of universal gravitation unifying Galileo's terrestrial mechanics and Kepler's laws of planetary motion, the so-called "Newtonian Synthesis" (Salmon, 1998, p. 85). This case is typically discussed as a example of theoretical convergence in the literature, most often specified in terms of *unification* or *consilience* (e.g., Friedman, 1983, Morrison, 2000). The classic reference, in this respect, is Willliam Whewell's analysis of Newton's achievement as a paradigmatic example of his notion of *Consilience of Inductions*, the second of his three confirmation criteria (i.e., novel predictions, consilience and coherence).[7] Whewell's consilience account of Newton's case has been variously interpreted, as we will see. To give some significant examples: Forster (1988) views it in terms of a common cause argument for realism, Harper (1989) as a "Natural Kind inference" providing evidential support, while Janssen (2002, pp. 488–89) understands it as a combination of "common-origin inferences" (*COIs*), that is, in his terminology, as a case of meta-*COI*.

As already apparent in the examples mentioned so far, *robustness*, *unification* and *consilience* are the key notions at stake when addressing CA cases. They form the core of the conceptual toolbox for discussing convergence arguments.[8] In the examples above, robustness was typically used in the first case, while the discussion of the second case was mostly conducted in terms of unification and consilience. In fact, as we will see, it is the kind of convergence argument that determines which notion is indeed significant and what role it actually plays in the argument. This will emerge more clearly by having a closer look at the tools available in the "convergence box". In particular, given the case studies considered in this paper, special attention will be devoted to robustness *vs* consilience arguments.

2.1. Robustness

The most discussed types of *CAs* are undoubtedly the so-called *robustness arguments*: that is, those assessment arguments that are based on the robustness of some scientific feature or result. In fact, as often noted in the literature, there are several notions of robustness, differing "both in their normative credentials and in the conditions that warrant their deployment" (Woodward, 2006, p. 219).

Historically, robustness was considered across models in its original version. This was in the celebrated article by Richard Levins on the strategy of model building in population biology (Levins, 1966). Levins notoriously made the following claims: first, that model building in the study of complex systems involves a necessary trade-off among generality, realism and precision; second, in order to solve this trade-off problem, "that the reliability of an inference is increased when it is the

[4] How to spell out the independence condition is not a simple issue and there is a lively debate in regard – see for example Stegenga (2012), Stegenga and Menon (2017), Schupbach (2018), Coko (2020b). Recent discussions of Bayesian accounts of the independence condition are provided, for example, in Stegenga and Menon (2017) and Landes (2020).

[5] For more details on this case, see section 3.

[6] See Schupbach (2018) and references therein. For further distinguo, an example is provided by Kuorikoski and Marchionni (2016, p. 238): they view Perrin's reasoning as an illustration of "the logic of triangulation", where triangulation is seen as a sort of robustness analysis but representing a different strategy with respect, for example, to what they call "Schupbach's eliminativist rationale".

[7] On the use of Whewell's account of Newton when discussing theoretical convergence, see for example Snyder (2008, p. 187). On Whewell's notion of consilience we'll focus later on, section 2.2.

[8] For the kinds of arguments we consider in this paper, the notion of coherence is of minor relevance.

joint inference of multiple models".[9] This latter was the claim about robustness: more precisely, in Levins' own words, "if these models, despite their different assumptions, lead to similar results, we have what we can call a robust theorem", whence his famous conclusion that "our truth is the intersection of independent lies" (1966, p. 423).

Since then, robustness has been the subject matter of a growing philosophical literature, especially after Wimsatt (1981)'s generalisation of Levin's ideas by developing a systematic account of "robustness analysis" for scientific reliability (e.g. Soler et al., 2012, and references therein). In Wimsatt's terms, "Things are robust if they are accessible (detectable, measurable, derivable definable, producible, or the like) in a variety of independent ways", and these "things" can be entities, properties, processes, results, or theorems (Wimsatt, 1994, p. 210). Thus, with respect to Levins' original formulation, Wimsatt extends robustness analysis to include a much larger variety of procedures, ranging from experimental manipulations, non-interventive observation or measurement to mathematical or logical derivation.[10]

However different, all these variants and uses of robustness have a "common theme" in Wimsatt's view: that is, distinguishing "that which is regarded as ontologically and epistemologically trustworthy and valuable from that which is unreliable, ungeneralizable, worthless, and fleeting" (Wimsatt, 1981, p. 63). How this claim about robustness could be effectively substantiated is a controversial issue in the literature, especially because of the diversity of the notions implied. Woodward (2006), for example, distinguishes four notions: inferential robustness (robustness as insensitivity of the results of inference to alternative specifications), derivational robustness, measurement robustness (robustness in agreement of measurement results), and causal robustness (robustness as a mark of causal or structural relationships). Calcott (2011), instead, identifies three kinds of robustness in Wimsatt's approach: robust theorems (theorems whose derivation can be supported in multiple ways), robust phenomena (phenomena which are reliably present in many different contexts), and robust detection (triangulation, multiple lines of evidence).[11] According to Eronen (2015, p. 3962) most of the discussion has focused on derivational robustness, while, he claims, there is "a more general form of robustness that is potentially more relevant for justifying inferences to what is real", that is, robustness as multiple accessibility.[12] Finally, Coko (2020b, 2022) distinguishes *multiple determination* – the epistemic strategy of using multiple, independent procedures to establish the same result - from variants of robustness analysis with which, as he argues, it is confused in the literature: while the first refers to the multiple, independent establishment of empirical claims about the world, the second refers to an analysis of some sort of invariance to change or perturbations.[13]

Whether these distinctions are indeed substantial and whether they cover the whole space of possibilities are debated issues.[14] In fact, there are two levels to consider in this debate. Let us assume in general that something is robust to the extent that it is obtained by many, different and independent means (e.g. Schupbach, 2018, p. 278).[15] Then, two levels of discussion can be distinguished in the literature. On the one side, the discussion regards the "operational level" of the concrete procedures (empirical or theoretical) employed for establishing the robustness. Examples are the analyses of robustness in terms of reliability (e.g. Basso, 2017), calibration (e.g. Bokulich, 2020), triangulation (e.g. Kuorikoski & Marchionni, 2016), and multiple determination (e.g., Coko, 2020b and 2022). On the other side, the discussion is focused on the "meta level" of the assessment strategies which are grounded on the robustness obtained at the first, operational level. Thus, to mention a concrete case, the issue at stake at this second level is not how Perrin succeeded in getting a robust result for Avogadro's number,[16] but rather the nature and import of his argument – based on the multiple determination of the empirical result – on behalf of the atomic hypothesis.

Here, we are specifically concerned with this second, "meta-level" type of analysis. Therefore, leaving aside questions regarding the operational level of the practices for getting a robust result, let us focus on the modalities of the robustness-based arguments for strengthening the trust in the viability of a theory (model, hypothesis). In this respect, we can distinguish three main types of strategies in the literature, depending on the starting points and the end goals of the convergence considered.

(1) A first type of strategy is the so-called robustness analysis developed in the framework of scientific model building (Levins, 1966 and 1993; Weisberg, 2006), based on *robustness across models*.[17] Starting with different idealised or approximate models, the aim is to arrive at a robust model or, possibly, at a true theoretical core. In Levins (1993, p. 554)'s terms, the strategy is as follows: given that, in science, "most of the models.. are partly true and partly false", we can "strengthen our confidence in the implications of some assumptions by using ensembles of models that share a common core of these assumptions but also differ as widely as possible in assumptions about other aspects."[18] In substance, the issue is how to deal with inevitably highly idealized models of complex systems, in order to determine which parts of these models make trustworthy predictions about their targets or can reliably be used in explanations (cfr. Weisberg, 2013, chap. 9).[19]

(2) A second strategy consists in increasing the confirmatory status of a given theory (model, hypothesis) by making it as robust as possible

[9] This is how Levins (1993, p. 550) characterises the claim about robustness in his 1966 paper.

[10] Wimsatt([1981] 2012, pp. 62–64) provides a list of all these procedures. See also Wimsatt (1994, p. 210).

[11] Calcott (2011, p. 284) underlines the epistemic character of this latter notion (in contrast to the ontological robustness of phenomena): "a claim about the world is robust when there are multiple, independent ways it can be detected or verified".

[12] According to Eronen (2015, p. 3962) these two notions correspond respectively to Calcott's notions of robust theorems and robust detection, and the second one is the kind of robustness that Wimsatt is mainly concerned with. Whether this corresponds indeed to the views of Wimsatt and Calcott is not a theme, here.

[13] This is not an exhaustive list of all the distinctions among different kinds of robustness to be found in the literature, of course. See also, for instance, Jones (2018, section 2.5).

[14] See for example Nederbragt (2012, p. 121); Coko (2022).

[15] We will not consider here more loose senses of robustness, making of it "a buzzword.. that can be applied to anything that exhibits strength of some sort", as Nickles (2012, p. 330) puts it.

[16] Woodward (2006, p. 234)'s discussion of Perrin's multiple determination of Avogardo's number as case of measurement robustness is a good example of such first-level type of analysis. This is well expressed by his comment that "It is common practice in many areas of science to take measurement robustness as grounds for increasing our confidence that the quantity has been measured accurately."

[17] See Coko (2022) for a detailed analysis of the accounts of Levins (1966, 1993), Wimsatt (1981) and Weisberg (2006, 2013), their rationale and their differences.

[18] More in detail, Levins's idea is that "the more the variable part spans the range of plausible assumptions, the more valid the claim that the conclusions shared by all of them depend on the constant part." Thus, "if we also have confidence that the constant part is true, then we have strong support for the claim that the conclusion is generally true. This gives robustness to the conclusions" (1993, p. 554).

[19] This kind of strategy has been applied especially to modeling complex systems, from biology to social and economical sciences. In the last decade, robustness analysis has been discussed also with respect to climate modeling (e.g., Lloyd, 2010, 2015; Parker, 2011, Weisberg, 2013, Vezér, 2017). More recently, robustness analysis has been extended to simulations in particles physics (e.g., Boge (in press)) and to evaluating cosmological modeling (e.g., Gueguen, 2020).

on the basis of *varied, independent evidence*.[20] In this perspective, robustness is usually (though not always) employed in a justification context, rather than in a discovery process. The rationale is that underlying the so-called "variety-of-evidence thesis" (varied evidence confirms more strongly than less varied evidence).[21] In the words of Hempel (1966, p. 34), to quote a classic historical reference in regard, "The confirmation of a hypothesis depends not only on the quantity of the favourable evidence available, but also on its variety: the greater the variety, the stronger the resulting support". In rough terms, the idea is that the chance of being simultaneous wrong in each of the different, independent evidential checks declines with increasing their numbers (e.g. Wimsatt, 1994, p. 210).[22] This kind of epistemic strategy has been much discussed in recent literature, especially in the framework of Bayesian approaches to confirmation.[23]

(3) Finally, a third type of strategy is that aiming at increasing the trust in a hypothesis or a theory on the grounds of the possibility of obtaining, on its basis, a robust derivation of a given result, which can be of empirical or theoretical nature. This third type of second-level modality, based on multiple determination (at the first, operational level),[24] is different from both the two second-level strategies mentioned above: on the one side, robustness is not considered across models (as in (1)); on the other side, it is a characteristic of the result obtained, not of the hypothesis/theory to be assessed (as in (2)). In other words, it is not just a case of varied evidence, since the evidence is the same one (the same result): what is varied, is the way of obtaining it, not the result itself.[25] In this case of assessment strategy, especially, the underlying rationale is often taken to be a *no-coincidence* (or *no-miracle*) argument, motivating (via an inference to the best explanation) the trust in the viability of the theoretical framework used for arriving at the robust result.[26] This view, however, has been criticised, either on the grounds of rejecting the epistemic import of this kind of no-coincidence argument,[27] or by proposing different accounts of robustness reasoning.[28]

[20] This seems to correspond to what Nederbragt (2012, p. 123) calls "multiple derivability", defined as "the strategy by which a theory is supported by the evidence obtained through two or more independent methods that differ in the background knowledge on which they are based".

[21] See Stegenga (2012, pp. 208-210) for a critical discussion of this rationale.

[22] In more explicit probabilistic terms, "If the probabilities of being correct, or of introducing error through an inference are both bounded between zero and one, then serial dependencies always reduce reliability and parallel redundancies always increase it" (Wimsatt, 1994, p. 210, fn.6). See also Woodward (2006, p. 234) for a similar rationale in the case of Perrin's measurement of Avogadro's number.

[23] Cfr. Landes (2020) for a recent analysis of this strategy in Bayesian terms. More generally, detailed discussions of the various aspects of the variety-of-evidence strategy are provided in the special issue by Fletcher et al. (2019). See also Vezér (2016, 2017) and references therein for analyses of variety-of-evidence reasoning as applied (and also critically discussed) in the context of climate model evaluation.

[24] Usually multiple determination is intended in reference to experimental procedures for determining experimental results (e.g., Coko, 2020b). In fact, it can be intended also in a more general way, including theoretical procedures as well as cases where the result is of theoretical nature.

[25] Of course, this does not mean that such strategies cannot be applied, and therefore also discussed, in combination (as it usually happens in most of the robustness literature).

[26] This view will be discussed in some detail in Section 3, in regard to the argument attributed to Perrin for the reality of the atomic hypothesis on the grounds of the multiple determination of Avogadro's number.

[27] With respect to Perrin's argument, for example, critical discussions of a no-coincidence interpretation based on an inference to best explanation are Cartwright (1983, pp. 82-84) and van Fraassen (2009).

[28] A different logic for robustness reasoning is discussed, for example, in Schupbach (2018).

2.2. Consilience

Beside robustness, convergence arguments are often analysed in terms of consilience and unification. Surely, there are close connections among these three notions. In particular, it is not always easy to disentangle one from each other when considering the role they play in specific convergence arguments. Consilience and unification, for example, are even treated as interchangeable in some literature.[29] Moreover, it is not rare to find discussions of consilience as a case of robustness reasoning (e.g., Wimsatt, 1981, p. 124).[30] Let us focus here on the notion of consilience and consider unification only insofar it is related to consilience in arguments for theory assessment.[31]

The term *consilience* is often used in today's philosophical and scientific literature in the loose sense of concordance or convergence *simpliciter*. When applied in more precise terms, the notion is taken to mean different things and with different epistemic import, depending on the context and case examined.[32]

In fact, even the original nineteenth century notion has not received an unanimous account in the scholarly literature. Whewell's own treatment of the notion has originated much discussion, giving rise to different interpretations (e.g. Hesse, 1968, 1971; Laudan, 1971, Forster, 1988, Harper, 1989; Snyder, 2006, 2008). Here, without entering into the detail of this interpretative issue, let us just focus on those relevant features of the original notion on which a convergence argument for theory assessment can be founded.

In his XIV aphorism among those "concerning science", Whewell (1840, p. 469) gives the following, famous characterization of the nature of consilience:

> "The Consilience of inductions takes place when an Induction, obtained from one class of facts, coincides with an Induction, obtained from another different class. This consilience is a test of the truth of the Theory in which it occurs."

Beside the multiplicity and independence of the evidence ("classes of facts altogether different"), what makes this "coincidence" or "agreement" a test of truth for hypotheses is also its *unexpectedness* – in Whewell (1840)'s own terms, an agreement "unforeseen and uncontemplated" (p. 65), "the unexpected coincidences of results drawn from distant parts of the subject" (p. 67). Note, in this regard, that what is unexpected – and therefore *surprising* – is the coincidence, not a new fact or prediction.

The epistemic role of this kind of surprise is well evident in Whewell's discussion of his most known example of consilience, that is Newton's Theory of Universal Gravitation.[33] The fact that Newton found that "the doctrine of the Attraction of the Sun varying according to the Inverse Square" of the distance, which explained Kepler's Third Law, explained also Kepler's First and Second Laws "although no connexion of these laws had been visible before", and that, again,"it appeared that the force of Universal Gravitation.. also accounted for the fact, apparently altogether dissimilar and remote, of the *Precession of*

[29] For example, Morrison (2000) often uses unification and consilience as interchangeable when discussing Whewell's account of Newton. See also, for a similar use, Friedman (1983, p. 242, fn. 14), and, more recently, Kao (2019, p. 3265).

[30] See also, for example, Nederbragt (2012, p. 123), describing Whewell's consilience of inductions as the oldest case of robustness strategies.

[31] The specific epistemic import of unification in convergence argument for theory assessment is analysed in another paper in preparation, in collaboration with Radin Dardashti and Richard Dawid.

[32] See for instance Fisch (1985) for an analysis of different types of consilience in contemporary's use of the notion. See also Coko (in press) for a discussion of the differences between Whewellian and today's notions of consilience.

[33] Another paradigmatic example is the Undulatory Theory of Light (Whewell, 1840, pp. 66-67).

the equinoxes" is, for Whewell, "a most striking and surprising coincidence, which gave to the theory a stamp of truth beyond the power of ingenuity to counterfeit" (1840, pp. 65-66).

Thus, the striking and surprising fact that the consilient theory can explain unrelated additional phenomena or laws is an essential part of the assessment argument (leading to an increase in the trust in the theory's truth). In other words, we can say that the coincidence or convergence must be surprising for consilience to function as an assessment argument.

Note that there are two levels at which the element of surprise is epistemically relevant, here: on the one side, the first level of the surprising fact of the convergence *per se*; on the other side, the meta level of the reasoning that – given the surprising convergence – it would be very surprising if the theory were false: in Whewell's words, "no accident could give rise to such an extraordinary coincidence" (1840, p. 65). It is only this second level of *reasoning from surprise* which is working in the no-miracle/no-coincidence argument often used for justifying the rationale of the robustness strategies (2) and (3) discussed in 2.1. There is no *surprise from unexpectedness* working at the first, operational level of robustness reasoning (cfr. 2.1): the fact that a result is obtained in multiple, different ways can be unlikely and asks for justification (for instance, by means of a no-coincidence IBE argument), but it is not unexpected *per se* – quite the opposite. In many cases, robustness is searched for, by varying circumstances, parameters, and so on. In this sense, if we want to transpose consilience in today's terms, there is an additional feature (the unexpectedness of the convergence) to be considered with respect to the reasoning from variety of evidence or multiple determination seen above.[34] However, this is often underestimated in current literature, where consilience is frequently identified with the convergence of multiple independent evidence "streams" or lines *tout court* (e.g. Forber & Griffith, 2011; Vezér, 2016; Currie, 2018a; Bokulich, 2020).[35]

In fact, beside the elements highlighted so far, understanding Whewell's consilience requires considering, in the framework of his particular theory of induction,[36] his notion of natural kind and common cause (e.g. Snyder, 2006; Coko, in press). As Snyder (2008, p. 187) puts it, "Consilience occurs when a theory brings together members of different kinds, showing that they belong to a more general classification. In the case of event kinds, individual types of events are members of the same kind when they share the same cause." This feature of "causal unification of different event or process kinds into more general kinds, in virtue of sharing a common cause" (*ibid.*) is precisely what has given rise, in the scholarly literature, to viewing Whewell's consilience in terms of *unification* and *common cause* (e.g. Forster, 1988, Harper, 1989, Janssen, 2002), as we have seen in the introductory part of section 2 (p. 4).

To sum up, how can we characterise a consilience argument for theory assessment in today's terms? As we have seen, depending on the context, interest and focus of the analysis, consilience is assimilated to different things in the literature: variety-of-evidence reasoning, multiple determination, natural kind inference, causal explanatory unification, a combination of "common-origin inferences". With respect to the convergence arguments seen in 2.1, however, there are two distinguishing features of consilience which emerge in the light of the genesis and development of the notion: i) the number of the different, independent evidence lines is not especially influential (already a small number of them are enough for boosting the trust in the theory's viability); ii) a distinct, fundamental role of surprise (corresponding to the unexpectedness of the convergence). In what follows, therefore, we will rely on these two elements for distinguishing consilience as a form of convergence argument.

3. *CA* arguments: Two representative cases

As history and scientific practice clearly show, convergence arguments for theory assessment are employed in a variety of cases of theory building. The question of interest, here, is how to understand the specificity of the epistemic import of these arguments – in particular, robustness and consilience arguments – from the viewpoint of theory assessment.

Let us address the question by considering two cases of *CAs*, the first one representative of a robustness argument, the second of a consilience argument: namely, a) the already mentioned case of Perrin's argument on behalf of the molecular hypothesis (hereafter, *Perrin's case*); b) the case of the convergence argument for boosting the trust in the "extra-dimension hypothesis" in the framework of early string theory (hereafter, *the extra-dimension case*). It is worth noting that Perrin's case, in addition to represent a robustness argument (as anticipated in section 2), is also commonly considered a typical example of empirical theory assessment, since the convergence is obtained in terms of measurements (however theory-mediated).[37] The extra-dimension case, on the contrary, will be shown to represent an instance of a consilience argument, as well as a case for non-empirical theory assessment: the argument is grounded on the convergence to a theoretical result (22 extra space dimensions), which is obtained in terms of theoretical procedures (although based on physical assumptions). Indeed, as we will see, this latter case is a clear example of how a non-empirical *CA* can be effective in motivating the acceptance of a very surprising hypothesis.[38]

(a) Perrin's case. The *CA* in question, in this case, is the argument attributed to Perrin on behalf of the atomic hypothesis on the grounds of the convergence of thirteen different procedures to the same result for the Avogadro number. As already mentioned in section 2, this argument has been seen in different ways in the scholarly literature: as an inference to the most probable cause (Cartwright, 1983), as a common cause argument (Salmon, 1984), as an instance of no-miracle argument (e.g. Chalmers, 2009, 2011, Psillos, 2011b), as a variety of robustness argument (e.g. Schupbach, 2018; Landes, 2020), as an example of calibration reasoning (Hudson, 2020a), and as a paradigmatic instance of multiple determination (Coko, 2020a, in press), to recall a number of stances.

Generally, these views are based on analyses of the concrete procedures followed by Perrin as well as on his own reflections. Here, since the paper's focus is on the "meta level" of convergence strate-

[34] Note that unexpectedness is not exactly the same thing as unlikeness, though the two notions are often related in the literature in terms of low priors in a probabilistic setting.

[35] A somewhat more restricted meaning is to be found in Psillos (2002, p. 615), where consilience is discussed in terms of competing explanations, by evaluating which explanation fits better with the background knowledge and should be, therefore, accepted as the best one. This meaning is taken over, for example, by Jones (2018, p. 88) and Linnemann (2020, p. 83). On the contrary, for a discussion of how consilience is different from IBE reasoning, see in particular Snyder (2008, p. 187).

[36] For Whewell, an induction or "colligation" of facts, involves a new idea or conception being "superinduced" on those facts in such a way that they are seen in a new way.

[37] There are many analyses of the procedures followed by Perrin in the literature, starting with his own writings. Historical reconstructions are, first of all, Brush (1968) and Nye (1972). For more recent, detailed analyses see, in particular, Bigg (2008), Chalmers (2009), Psillos (2011b), Hudson (2020a), Coko (2020a), Smith and Seth (2020), Demopoulos (2022).

[38] A discussion of the case of the extra-dimension conjecture in early string theory is provided in Castellani (2012). Castellani (2019) focuses again on this story, reconstructing it as an example of scientific methodology based on a convergence argument in non-empirical theory assessment. Linnemann (2020) uses Castellani (2019)'s account of the different paths to arrive at the extra-dimension conjecture as an example of non-empirical robustness argument. In what follows, we will analyse in more detail the kind of convergence argument represented by this case, showing that it is more appropriate to see it in terms of consilience rather than in terms of robustness.

gies for theory assessment, I will not be concerned with the details of Perrin's measurement procedures as rather with the following question: whether, in Perrin's case, there is a distinctive, epistemic import due to the fact of the convergence with respect to mere empirical confirmation (however strong or "robust"), and, if this is the case, in what this additional epistemic feature actually consists.

The *CA* attributed to Perrin is basically grounded on a number of Perrin's famous statements. A most quoted one is the following conclusion of his review of the various phenomena yielding concordant values for Avogadro's constant in his book *Les Atomes* (Perrin, [1913] 1916, pp. 206-7):

> Our wonder is aroused at the very remarkable agreement found between values derived from the consideration of such widely different phenomena. Seeing that not only is the same magnitude obtained by each method when the conditions under which it is applied are varied as much as possible, but that the numbers thus established also agree among themselves, without discrepancy, for all the methods employed, the real existence of the molecule is given a probability bordering on certainty.

This passage is representative of many similar reflections to be found in Perrin's writings.[39] In the literature, these remarks are usually taken to indicate that, according to Perrin, the trust in the truth of the molecular (atomic) hypothesis is boosted on the grounds of obtaining, on its basis, a robust result for the number of molecules in a mole (whatever other assumptions are used in the different procedures for arriving at the value of Avogadro's number).[40] In other words, in terms of the distinctions introduced in section 2.1, the argument attributed to Perrin can be seen as a case of robustness *CA* corresponding to the third type of convergence strategy discussed.

Now, whether this sort of argument has effectively played a significant role in viewing Perrin's contribution as conclusive for establishing the existence of atoms is a debated issue, from both a historical and an epistemic point of view.[41] From this latter point of view, in particular, much of the discussion has focused on the presumed rationale of the argument. The key question regards the distinctive epistemic role to be attributed to the convergence of the many, independent ways to obtain Avogadro's number in assessing the atomic hypothesis, to which we will turn in some detail in the next section.

(b) The extra-dimension case. The context for discussing this case is the so-called Early String Theory (EST): that is, the first developments of string theory from the 1968 formulation by Gabriele Veneziano of his famous scattering amplitude to the *first string revolution* in 1984.[42] In the framework of this "founding era" of string theory, the extra-dimension conjecture emerged in the first phase, characterised by the developments of the *dual theory of strong interactions* in the years 1968-1973.[43]

This EST initial phase was originally aimed at finding a viable theory of hadrons in the framework of the so-called analytic S-matrix (or S-matrix theory) developed in the early Sixties.[44] Its programme was to determine the relevant observable physical quantities, i.e. the scattering amplitudes, only on the basis of some general principles such as *unitarity*, *analiticity* and *crossing symmetry* and a minimal number of additional assumptions, among which the so-called *duality principle*.[45]

In this framework, the problem of finding a scattering amplitude obeying also the duality principle was brilliantly solved by Veneziano for the case of four mesons. This ground-breaking result, universally recognised as the starting point of string theory, immediately gave rise to a period of intense theoretical activity aimed at extending Veneziano's amplitude: from the first two models for the scattering of N particles – the generalised Veneziano model, known as the *Dual Resonance Model* (DRM), and the *Shapiro-Virasoro Model*[46] – to all the subsequent endeavours to extend, complete and refine the theoretical framework, including its string interpretation and the addition of fermions (see Cappelli et al., 2012, Part III).

Two particularly significant conjectures were introduced in this process. First, the string conjecture in 1969: in independent attempts to gain a deeper understanding of the physics described by dual amplitudes, Nambu, Nielsen and Susskind each arrived at the conjecture that the underlying dynamics of the dual resonance model was that of a quantum-relativistic oscillating string.[47] Second, the conjecture or "discovery" of extra spacetime dimensions: independent developments of the dual theory led to the critical value $d = 26$ for the spacetime dimension (the *critical dimension*), reducing to the value $d = 10$ when including fermions.

In what follows, we briefly illustrate the three independent theoretical processes leading – by *surprisingly* converging to the same *surprising* result ($d = 26$) – the research community to accept the critical-dimension conjecture, however bold and apparently unphysical.[48]

• *Three ways to the critical dimension*

In the framework of the theoretical endeavours to extend the original dual theory in order to overcome its initial limitations and problems, the critical dimension conjecture first emerged in the context of two independent programmes: 1) the "unitarization programme", in the context of which Claud Lovelace arrived at the conjecture $d = 26$ while addressing a problematic singularity case arising in the construction of the nonplanar one-loop amplitude; 2) the "ghost elimination programme", where the critical value $d = 26$ for the spacetime dimension issued from studying the spectrum of states of the Dual Resonance Model. In some more details:

1. *Lovelace's result.* The original dual amplitudes didn't respect the S-matrix unitarity condition. To go beyond the initial narrow-resonance approximation, the "unitarization programme" substantiated in generalising the initial amplitudes, considered as the lowest order or tree diagrams of a perturbative expansion, to include loops. As a first step

[39] Psillos (2011a, 2011b, 2014) discusses many samples of Perrin's passages. More recently, see for example Hudson (2020a), Coko (2020a), Smith and Seth (2020, chap. 6), and Demopoulos (2022, chap. 2).

[40] Bigg (2008, p. 316) puts it nicely: "Perrin's concordance argument was simple and effective: by giving an impressive list of theories and experiments, by himself or others, that all led to comparable values for N, he made atoms the meaningful link between all these unrelated phenomena."

[41] On how Perrin's contribution was effectively received in his times, classic references are Brush (1968) and Nye (1972). More recent analyses are to be found, for example, in Psillos (2011b), (2014); Coko (2020a); Hudson (2020a); Smith and Seth (2020). See also Demopoulos (2022, 2.1), for a careful analysis of the reasons why Perrin's contribution was accorded the status of a turning point in the assessment of the molecular hypothesis.

[42] For details, see Cappelli et al. (2012), providing a thorough historical reconstruction of the early developments of string theory.

[43] On the distinction between two phases of EST - a first phase (1968-1973) ending with the falsification of EST as a theory of strong interactions, and a second phase (1974-1984), where the theory was re-interpreted as a unified quantum theory of all fundamental interactions, see Castellani (2012, section 1). For detailed descriptions of these developments, see Cappelli et al. (2012, Part 1).

[44] On the S-matrix programme pursued by Chew and his collaborators, see in particular Cushing (1990), and Cappelli et al. (2012, Part II).

[45] The meaning of this duality (also known as *DHS duality* after the physicists Dolen, Horn and Schmid, who introduced it in 1957 on the grounds on experimental data) was that the contributions from resonance intermediate states and from particle exchange each formed a complete representation of the scattering process (so that they should not be added to one another in order to obtain the total amplitude). For more details on this duality and its relevance see Castellani (2012, p. 68), and Cappelli et al. (2012, Part II, 5.4.3).

[46] These two models were later understood as describing open and closed strings, respectively.

[47] For details, see Castellani (2012, pp. 72-73); Cappelli et al. (2012, Part IV).

[48] For a detailed description of the independent ways to arrive at the critical dimension, see Castellani (2019, 3.1).

for restoring unitarity, one-loop diagrams were constructed, and in this building process the calculation of a nonplanar loop diagram led Lovelace, in order to solve a singularity problem emerged in the process, to the 1971 conjecture of the value $d = 26$ for the spacetime dimension.[49]

2. *The "no ghost" result.* In the endeavours for generalising Veneziano's amplitude to the scattering of an arbitrary number N of scalar particles, a serious problem was represented by the presence of negative-norm states ("ghosts") in the state spectrum of the model.[50] These states, leading to unphysical negative probabilities, had to be eliminated from the theory. In this "ghost elimination" programme, a decisive step was the 1971 construction by Del Giudice, Di Vecchia and Fubini of an infinite set of positive-norm states (the so-called DDF states), which were found to span the whole space of physical states if the spacetime dimension d was equal to 26. Soon after, the proof of the so-called No-Ghost Theorem, establishing that the Dual Resonance Model has no ghosts if $d \leq 26$, was achieved by Brower, and independently by Goddard and Thorn.[51]

A spacetime of 26 dimensions was not easy to accept.[52] While initially almost nobody had taken Lovelace's conjecture seriously, after the proof of the No-Ghost Theorem the attitude changed and the extra dimensions started to be gradually accepted in the dual model community.[53] A further decisive support to the conjecture came from the third theoretical process leading, independently from the previous two ways, to the same "critical" value $d = 26$ for the spacetime dimension: the 1973 work of Goddard, Goldstone, Rebbi and Thorn (GGRT) on the quantisation of the string action.

3. *The GGRT result.* After the 1969 string conjecture and the immediately successive studies of a Lagrangian action for the string,[54] the quantisation of the string action by Goddard, Goldstone, Rebbi and Thorn was a decisive step for the string interpretation of the dual resonance model to be fully accepted. In the resulting quantized theory, all what had been previously obtained by proceeding according to a bottom-up approach and following different paths could now be derived in a more clear and unitary way. In particular, the critical dimension was obtained as a condition for the Lorentz invariance of the canonical quantisation of the string in the light-cone gauge: only for $d = 26$ the quantisation procedure was Lorentz invariant.[55]

Of course, the story of the critical dimension goes further, and other decisive support to this conjecture came from successive developments of string theory, especially after it was re-interpreted as a unified quantum theory of all fundamental interactions including gravity.[56] But let's stop at this point and turn to consider the rationale of the convergence arguments operating in the two representative cases described so far.

4. Conclusion: The interpretative issue

The two cases of *CAs* considered in the previous section are surely very different. They represent distinct types of convergence arguments (robustness *vs* consilience) and, in addition, different cases of theory assessment (empirical *vs* non empirical). In both cases, however, the trust in the theory or hypothesis involved is undoubtedly boosted on the basis of the convergence. In which way, exactly? This section is devoted to address this point by examining the kind of epistemic strategy at work in each case, and the related interpretative issue.

(a) Perrin's case. As already said, different interpretations of the rationale behind Perrin's reasoning have been proposed and discussed in the literature.[57] Nonetheless, there is a substantial agreement on the fact that the "miracle of concordances" (Psillos, 2011b, p. 360) has played a significant role in boosting the trust in the atomic hypothesis. Salmon (1998) notoriously comments on Perrin's multiple determination of Avogadro by noting that "such agreement would be miraculous il matter were not composed of molecules and atoms" (p. 82). In a similar vein, Chalmers (2009, p. 243) remarks: "The concordance of a variety of indisputable evidence with the predictions of the kinetic theory amounted to a powerful argument from coincidence. How could the theory get things so right if it were not at least roughly true?".

In fact, the epistemic relevance of an "argument from coincidence" in this case has been understood in a number of different ways over the years. Cartwright (1983, p. 82), for example, argues that, while for many it is "a paradigm of inference to the best explanation", what Perrin really makes is "a more restricted inference – an inference to the most probable cause".[58] Coming to the current stage of this long-standing debate, the details of which have been thoroughly analysed in recent literature,[59] a new, more sophisticated way of seeing Perrin's reasoning as a no-coincidence argument is offered by Coko (2020a, 2020b) in terms of his "multi-dimensional approach": the epistemic force of the argument, according to Coko, depends on the modality of the concurrence of the several elements ("dimensions") of multiple determination, such as the independence, reliability and number of the converging procedures. On this view, the epistemic import of a *CA* has to be analysed

[49] In four spacetime dimensions, the amplitude had a singularity (a 'branch cut') in a certain channel, incompatible with unitarity. Lovelace realised that the singularity could be turned into a *pole*, and thus interpreted as due to the propagation of a new intermediate particle, if the value of the spacetime dimension was $d = 26$. This pole, Lovelace conjectured to be the Pomeron, the particle that was later understood as the graviton. See Lovelace's own account of his discovery in Cappelli et al. (2012, Chapter 15). For more details, see Cappelli et al. (2012, Section 10.2.3).

[50] Note that this is a different meaning of the term "ghost" with respect to how it is commonly used in quantum field theory (i.e., to indicate the unphysical fields associated with gauge invariance in functional approaches to field theory quantisation).

[51] By essentially same argument as in the case of the DRM, it was also proved that Neveu-Schwarz dual model has no ghosts if $d \leq 10$, thus confirming the critical dimension as $d = 10$ in the case including fermions. A detailed description of the No-Ghost result can be found, in particular, in Goddard's contribution to Cappelli et al. (2012, Chapter 20).

[52] In a recollection paper on his contribution to the dual theory, Lovelace describes the first reactions to his conjecture as follows: "I gave a seminar... which was attended by some powerful people as well as the Dual Model group. Treating the result as a joke, I said I had bootstrapped the dimension of spacetime but the result was slightly too big. Everyone laughed" (Castellani (2019), p. 179; Cappelli et al. (2012), Chapter 15, p. 228).

[53] A good example is given in the following quote by Goddard (Cappelli et al. (2012, Chapter 20, p. 285): "The validity of the No-Ghost Theorem had a profound effect on me. It seemed clear that this result was quite a deep mathematical statement..., but also that no pure mathematician would have written it down. It had been conjectured by theoretical physicists because it was a necessary condition for a mathematical model of particle physics not to be inconsistent with physical principles.... I could not help thinking that, in some sense, there would be no reason for this striking result to exist unless the dual model had something to do with physics, though not necessarily in the physical context in which it had been born."

[54] Nambu (and then Goto) proposed the Lagrangian action for the string formulated in terms of the area of the surface swept out by a one-dimensional extended object moving in spacetime, in analogy with the formulation of the action of a point particle in terms of the length of its trajectory.

[55] Details on this point, and in general on the quantisation of the hadronic string, are provided by Di Vecchia and Goddard in their contributions to Cappelli et al. (2012, Chapter 11, 11.8 and Chapter 20, 20.7), respectively.

[56] See Castellani (2019, pp. 181-83); Cappelli et al. (2012, Part VI).

[57] For a detailed reconstruction of different views on Perrin's argument, see for example Coko (2020a).

[58] More precisely, her argument is that "Coincidence enters Perrin's argument, but not in a way that supports inference to the best explanation in general. [...] Coincidence will not help with laws. We have no ground for inferring from any phenomenological law that an explanatory law must be just so; multiplying cases cannot help." (1983, p 84).

[59] See, for example, Coko (2020a), Hudson (2020a), Smith and Seth (2020), and Demopoulos (2022, 2.1).

case by case, by looking at how well the different dimensions are instantiated.[60]

An alternative point of view is provided by Dawid (2021), who argues that, beside the implausibility of the coincidence scenario, two meta-empirical criteria are needed "for making a convincing case for atomism based on Perrin's results": namely, the absence of no non-atomist explanation other than mere coincidence, and the unlikeness of the existence of unconceived alternative explanations. Also Smith and Seth (2020) do not endorse a no-coincidence account based on IBE reasoning, although from a different perspective.[61] More precisely, in reflecting on the evidential significance (for the reality of molecules) of Perrin's "converging theory-mediated measurements", they propose to understand the force of the evidence provided by the convergence by construing it "as a form of same-effect-same-cause reasoning – specifically as same-magnitude-same-quantity-being-measured reasoning" (p. 310).[62]

Finally, a further, different way of intending the import of the convergence in Perrin's case is offered by Demopoulos (2022, chap. 2). Demopoulos rejects "an account of Perrin's success that is based on the hypothetico-deductive method or the method of inference to the best explanation", while, at the same time, maintaining a realist understanding of the argument.[63] Perrin's argument, for Demopoulos, is indeed an argument for molecular reality, but "it has a subtlety that is easily missed" (p. 78). In order to show how it concretely works, Demopoulos provides a careful reconstruction of Perrin's argument articulated in five stages. As regards specifically the concordance, its epistemic role enters at the fourth stage: that is, the stage which "consists in recounting the support that the connecting link [for the empirical determination of a host of molecular parameters] receives from the remarkable uniformity and concordance of the determination of parameter values to which it leads with various other determinations of these parameter values" (p. 92). Then, the final (fifth) stage "infers from what the earlier stages have revealed the explanation of Brownian motion in terms of the molecular hypothesis" (ibid.).

To sum up, we can say that there is a shared agreement, in these representative positions, that the concordance of the various determinations of Avogadro's constant plays a distinctive, additional epistemic role in Perrin's reasoning besides mere empirical confirmation. How, then, this role is precisely specified – whether in terms of an inference to the best explanation, in terms of an inference to the most probable cause, in terms of a same-effect-same-cause inference or in terms of meta-empirical reasoning – significantly depends on the interpretative stance adopted, as we have seen.

(b) The extra-dimension case. As described in the previous section, the critical-dimension conjecture emerged from endeavours to extend the original dual theory and thus overcome its initial limitations and problems. These endeavours were mostly of theoretical nature, but justified or motivated on the grounds of the *physics* studied – that is, let us stress, on the grounds of assumptions and constraints of *both* phenomenological and theoretical nature.[64] In this theory-building process, characterised by a close interplay of mathematically driven creativity and physical constraints, the fact that the value $d = 26$ for the spacetime dimension was obtained in three different, independent ways surely was an influential reason for taking it seriously. In fact, already after the second result (i.e., the no-ghost one), the initial skeptical attitude started to change.

Now, a first question is whether this fact can be considered a sufficient basis for a convergence argument on behalf of the viability of the extra-dimension hypothesis. The real independence of the three ways leading to the critical dimension could be questioned, for example. But, analogously, one could question the independence of the different lines of evidence in many other (empirical) convergence cases, including Perrin's one, as has been done by some authors.[65] Assuming, to the contrary, that the convergence of the different paths to the same surprising numerical result $d = 26$ provides a legitimate convergence argument for boosting the trust in the extra-dimension conjecture, the question becomes: what are the distinctive features of the argument doing the epistemic work in this case?

First of all, an analysis in terms of robustness does not seem to be appropriate, here. To start with, the number of independent ways of arriving at the result is very low (when compared to the thirteen ways of Perrin's case, for example). Beside, the convergence is not searched for. Quite the opposite: it is completely unexpected. Moreover, the result itself is very surprising.

As underlined in section 2.2, these features – a (small) number of concordant procedures, the unexpectedness of the convergence and the role of surprise – can be taken as the distinguishing characteristics of consilience as a convergence argument for theory assessment. The fact that we are dealing with a non-empirical case – in the sense that the convergence is to a theoretical result, obtained on the grounds of theoretical procedures (though *physically* motivated, it is worth recalling) – is not relevant from the viewpoint of the consilience structure of the argument.

The extra-dimension case can thus be seen as a particular instance of consilience. As highlighted in the previous section, the element of surprise plays a distinctive epistemic role in boosting the trust in an hypothesis in consilience cases. Actually, there are two kinds of surprising facts in this specific case: a) the convergence of the different, independent paths to the same numerical result $d = 26$ for the spacetime dimension; b) the result itself, which is undoubtedly very surprising. Correspondingly, the surprise factor has a double role here: first, by motivating a no-coincidence argument on behalf of the extra-dimension conjecture; second, by providing further support to the force of the argument.

Of course, this is a very particular case of theory assessment, where no empirical evidence is available. In this sense, it is naturally different from other examples of consilience, such as the case of Newton's unification discussed by Whewell (see section 2). However, it is worth underlining, the specific epistemic role of surprise in the consilience argument is independent of the empirical or non-empirical nature of the case considered. In other words, the assessment strategy based on consilience can be successful or defective in both cases, depending on how the epistemic import of consilience is interpreted. In current literature, there is a growing attention to the epistemic role of surprise in scientific pratice (e.g., Currie, 2018b; French & Murphy, 2023). Without entering in the details of this discussion,[66] what is of interest to underline, here, is that we can draw a similar conclusion for this kind of consilience argument as in the previous robustness case: the evaluation of the assessment strategy in such cases significantly depends on how the convergence argument is interpreted. And this is independent of whether we are dealing with an empirical case of convergence argument, as in the Perrin's case, or with a non-empirical case, as for the extra-dimension conjecture.

[60] See also Coko (2020b, 2022)'s distinction between multiple determination (based on a no-coincidence argument) and "robustness analysis" as developed in the different accounts by Levins (1966, 1993), Wimsatt (1981, 2007) and Weisberg (2006), based on an underlying rationale of elimination.

[61] In particular, they critically discuss realist no-coincidence accounts such as those of Chalmers (2009) or Psillos (2011b, 2014).

[62] The authors explicitly refer, here, to Newton's first Rule of Reasoning.

[63] In this respect Demopoulos strongly disagrees with van Fraassen (2009), "who interprets Perrin as having only been concerned to show the empirical determinability of various parameters of the molecular-kinetic theory, so that the theory could be seen to be empirically grounded" (p. 78, fn. 19).

[64] For details, see for example Castellani (2012, 4.3).

[65] See section 2, fn.4.

[66] For what concerns specifically consilience and unification, the epistemic role of surprise is the subject matter of a paper in preparation with Radin Dardashti and Richard Dawid.

CRediT authorship contribution statement

Elena Castellani: Conceptualization, Writing – original draft, Writing – review & editing.

Data availability

No data was used for the research described in the article.

Acknowledgements

Many thanks for helpful comments to the audiences of both the 2019 Berlin workshop "Non-Empirical Physics from a Historical Perspective" and the 2019 Stockholm workshop on "Non-Empirical Theory Assessment", where a previous version of this paper was presented. I am especially grateful to Radin Dardashti and Richard Dawid for many invaluable discussions over the last few years, and to Klodian Coko for his generous feedback. A special thank to Casey McCoy for his support in making the paper reach its final stage. Thanks also to two anonymous referees for their comments and suggestions. This work was supported by the Italian Ministry of Education, University and Research through the PRIN 2017 Program The Manifest Image and the Scientific Image (Prot.2017ZNWW7F–004).

References

Basso, A. (2017). The appeal to measurement in experimental practice. *Studies in History and Philosophy of Science A*, *65–66*, 57–66.
Bigg, C. (2008). Evident atoms: Visuality in jean perrin's Brownian motion research. *Studies in History and Philosophy of Science*, *39*, 312–322.
Boge, F.J. (in press). Why Trust a Simulation? Models, Parameters, and Robustness in Simulation-Infected Experiments the British Journal for the Philosophy of Science.
Bokulich, A. (2020). Calibration, coherence, and consilience in radiometric measures of geologic time. *Philosophy of science*, *87*(3), 425–456.
Brush, S. G. (1968). A history of random processes: I. Brownian movement from brown to perrin. *Archive for History of Exact Sciences*, *5*(1), 1–36.
Calcott, B. (2011). Wimsatt and the robustness family: Review of wimsatt's re- engineering philosophy for limited beings. *Biology & Philosophy*, *26*, 281–293.
Cappelli, A., Castellani, E., Colomo, F., & Di Vecchia, P. (Eds.). (2012). *The birth of string theory*. Cambridge: Cambridge University Press.
Cartwright, N. (1983). *How the laws of physics Lie*. Oxford: The Clarendon Press.
Castellani, E. (2012). Early string theory as a challenging case study for philosophers. In A. Cappelli, et al. (Eds.), *The Birth of String Theory* (pp. 71–89). Cambridge: Cambridge University Press.
Castellani, E. (2019). Scientific methodology: A view from early string theory. In R. Dardashti, R. Dawid, & K. Thébault (Eds.), *Why trust a theory?: Epistemology of fundamental physics* (pp. 173–183). Cambridge: Cambridge University Press.
Chalmers, A. (2009). *The scientist's atom and the philosopher's stone*. Dordrecht: Springer.
Chalmers, A. (2011). Drawing philosophical lessons from Perrin's experiments on Brownian motion: A response to van Fraassen. *The British Journal for the Philosophy of Science*, *62*(4), 711–732.
Cushing, J. T. (1990). *Theory construction and selection in modern physics: The s-matrix*. Cambridge: Cambridge University Press.
Coko, K. (2020a). Jean Perrin and the PhilosophersÕ stories: The role of multiple determination in determining Avogadro's number. *HOPOS*, *11*, 143–193.
Coko, K. (2020b). The multiple dimensions of multiple determination. *Perspectives on Science*, *28*(4), 505–541.
Coko, K. (in press). Hypothesis and Consilience in the Long Nineteenth Century in: E. Crull, & E. Peterson (Eds.), History and Philosophy of Modern Science: 1750-1900. Bloomsbury Press.
Coko, K. (2022). Variants of Robustness and Multiple Determination. Unpublished draft.
Currie, A. (2018a). *Rock, Bone, and Ruin: An Optimist's Guide to the Historical Sciences*. Cambridge, MA: The MIT Press.
Currie, A. (2018b). The argument from surprise. *Canadian Journal of Philosophy*, *48*(5), 639–661.
Dardashti, R., Dawid, R., & Thebault, K. (Eds.). (2019). *Why trust a theory? Epistemology of fundamental physics*. Cambridge: Cambridge University Press.
Dawid, R. (2013). *String theory and the scientific method*. Cambridge: 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*, *90*, 50–60.
Demopoulos, W. (2022). *On theories. logical empiricism and the methodology of modern physics*. Cambridge, MA: Harvard University Press.
Eronen, M. I. (2015). Robustness and reality. *Synthese*, *192*, 3961–3977.
Fletcher, S. C., Landes, J., & Poellinger, R. (2019). Evidence amalgamation in the sciences: An introduction. *Synthese*, *196*, 3163–3188.
Fisch, M. (1985). Whewell's consilience of inductions: An evaluation. *Philosophy of Science*, *52*(2), 239–255.
Forber, P., & Griffith, E. (2011). Historical reconstruction: Gaining epistemic access to the deep past. *Philosophy, Theory, and Practice in Biology*, *3*(3), 203–222.
Forster, M. R. (1988). Unification, explanation, and the composition of causes in Newtonian mechanics. *Studies in History and Philosophy of Science*, *19*(1), 55–101.
French, S., & Murphy, A. (2023). The value of surprise in science. *Erkenntnis*, *88*, 1447–1466.
Friedman, M. (1983). *Foundations of space-time theories*. Princeton: Princeton University Press.
Gueguen, M. (2020). On robustness in cosmological simulations. *Philosophy of Science*, *87*(5), 1197–1208.
Harper, W. (1989). Consilience and natural kind reasoning. In J. R. Brown, & J. Mittelstrass (Eds.), *An intimate relation* (pp. 115–152). Dordrecht: D. Reidel.
Hempel, C. (1966). *Philosophy of natural science*. Englewood Cliffs, N.J.: Prentice-Hall.
Hesse, M. B. (1968). Consilience of inductions. In I. Lakatos (Ed.), *The problem of inductive logic* (pp. 232–246). Amsterdam: North-Holland.
Hesse, M. B. (1971). Whewell's consilience of inductions and predictions. *The Monist*, *55*, 520–524.
Hossenfelder, S. (2018). *Lost in math. how beauty leads physics astray*. New York: Basic Books.
Hudson, R. (2020a). What was Perrin really doing in his proof of the reality of atoms? *HOPOS*, *10*, 194–218.
Hudson, R. (2020b). The reality of Jean Perrin's atoms and molecules. *The British Journal for the Philosophy of Science*, *71*, 33–58.
Janssen, M. (2002). COI stories: Explanation and evidence in the history of science. *Perspectives on Science*, *10*(4), 457–522.
Jones, N. (2018). Inference to the more robust explanation. *The British Journal for the Philosophy of Science*, *69*(1), 75–102.
Kao, M. (2019). Unification beyond justification: A strategy for theory development. *Synthese*, *196*, 3263–3278.
Kuhn, T. (1970). Reflections on my critics. In I. Lakatos, & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 231–238). Cambridge: Cambridge University Press.
Kuorikoski, J., & Marchionni, C. (2016). Evidential diversity and the triangulation of phenomena. *Philosophy of Science*, *83*, 227–247.
Landes, J. (2020). Variety of evidence. *Erkenntnis*, *85*, 183–223.
Laudan, L. (1971). William whewell on the consilience of inductions. *The Monist*, *55*(3), 368–391.
Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, *54*, 421–431.
Levins, R. (1993). A response to Orzack and Sober: Formal analysis and the fluidity of science. *The Quarterly Review of Biology*, *68*(4), 547–555.
Linnemann, N. S. (2020). Non-empirical robustness arguments in quantum gravity. *Studies in History and Philosophy of Modern Physics*, *72*, 70–86.
Lloyd, E. A. (2010). Confirmation and robustness of climate models. *Philosophy of Science*, *77*(5), 971–984.
Lloyd, E. A. (2015). Model robustness as a confirmatory virtue: The case of climate science. *Studies in History and Philosophy of Science*, *49*, 58–68.
Morrison, M. (2000). *Unifying scientific theories*. New York: Cambridge University Press.
Nederbragt, H. (2012). Multiple derivability and the reliability and stabilization of theories. In L. Soler, E. Trizio, T. Nickles, & W. Wimsatt (Eds.), *Characterizing the robustness of science: After the practice turn in the philosophy of science* (pp. 121–145). Dordrecht: Springer.
Nickles, T. (2012). Dynamic robustness and design in nature and artifact. In L. Soler, E. Trizio, T. Nickles, & W. Wimsatt (Eds.), *Characterizing the robustness of science: After the practice turn in the philosophy of science* (pp. 329–360). Dordrecht: Springer.
Nye, M. J. (1972). *Molecular reality: A perspective on the scientific work of Jean Perrin*. London: MacDonald.
Parker, W. S. (2011). When climate models agree: The significance of robust model predictions. *Philosophy of Science*, *78*(4), 579–600.
Penrose, R. (2016). *Fashion, faith, and fantasy in the new physics of the universe*. Princeton: Princeton University Press.
Perrin, J. ([1913] 1916). *Atoms*. New York: D. van Nostrand.
Psillos, S. (2002). Simply the best: A case for abduction. In A. C. Kakas, & F. Sadri (Eds.), *Computational logic: Logic programming and beyond* (pp. 605–625). Berlin: Springer.
Psillos, S. (2011a). Making contact with molecules: On Perrin and achinstein. In G. J. Morgan (Ed.), *Philosophy of science matters: The philosophy of Peter achinstein* (pp. 177–190). Oxford: Oxford University Press.
Psillos, S. (2011b). Moving molecules above the scientific horizon: On Perrin's case for realism. *Journal for General Philosophy of Science*, *42*, 339–363.
Psillos, S. (2014). The view from within and the view from above: Looking at van Fraassen's Perrin. In W. J. Gonzalez (Ed.), *Bas van Fraassen's approach to representation and models in science* (pp. 143–166). Dordrecht: Springer.
Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
Salmon, W. (1998). *Causality and explanation*. New York: Oxford University Press.
Schupbach, J. (2018). Robustness analysis as explanatory reasoning. *The British Journal for the Philosophy of Science*, *69*, 275–300.

Smith, G. E., & Seth, R. (2020). *Brownian motion and molecular reality*. New York: Oxford University Press.

Snyder, L. J. (2006). *Reforming philosophy: A Victorian debate on science and society*. Chicago: University of Chicago Press.

Snyder, L. J. (2008). The whole box of tools: William Whewell and the logic of induction. In J. Woods, & D. Gabbay (Eds.), *The handbook of the history of logic (Vol. IV)* (pp. 163–228). Dordrecht: Kluwer.

Soler, L., Trizio, E, Nickles, T., & Wimsatt, W. C. (Eds.). (2012). *Characterizing the robustness of science: After the practice turn in the philosophy of science*. Dordrecht: Springer.

Stegenga, J. (2012). Rerum concordia discors: Robustness and discordant multimodal evidence. In L. Soler, E. Trizio, T. Nickles, & W. Wimsatt (Eds.), *Characterizing the robustness of science: After the practice turn in philosophy of science* (pp. 207–226). Dordrecht: Springer.

Stegenga, J., & Menon, T. (2017). Robustness and independent evidence. *Philosophy of Science*, *84*(3), 414–435.

van Dongen, J. (2021). String theory, Einstein, and the identity of physics: Theory assessment in absence of the empirical. *Studies in History and Philosophy of Science*, *89*, 164–176.

van Fraassen, B. C. (2009). The perils of Perrin, in the hands of philosophers. *Philosophical Studies*, *143*, 5–24.

Vezér, M. A. (2016). Computer models and the evidence of anthropogenic climate change: An epistemology of variety-of-evidence inferences and robustness analysis. *Studies in History and Philosophy of Science*, *56*, 95–102.

Vezér, M. A. (2017). Variety-of-evidence reasoning about the distant past. *European Journal for Philosophy of Science*, *7*(2), 257–265.

Weisberg, M. (2006). Robustness analysis. *Philosophy of Science*, *73*, 730–742.

Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. New York: Oxford University Press.

Whewell, W. ([1840], 1847). *The philosophy of the inductive sciences, founded upon their history* (2d ed.). London: John W. Parker.

Wimsatt, W. C. (1981). Robustness, reliability, and overdetermination. In M. B. Brewer, & B. E. Collins (Eds.), *Scientific inquiry and the social sciences* (pp. 124–163). San Francisco: Jossey-Bass. Reprinted in L. Soler, E. Trizio, T. Nickles, W. Wimsatt (Eds.), *Characterizing the Robustness of Science: After the Practice Turn in the Philosophy of Science*, Dordrecht: Springer (pp. 61–87).

Wimsatt, W. C. (1994). The ontology of complex systems: Levels of organization, perspectives, and causal thickets. In *Canadian journal of philosophy*. Reprinted in*Re-engineering philosophy for limited beings. piecewise approximations to reality*. Cambridge, MA: Harvard University Press (pp. 193–240).

Woodward, J. (2006). Some varieties of robustness. *Journal of Economic Methodology*, *13*(2), 219–240.