Theory Success: Some Evaluative Clues

María Caamaño-Alegre
Universidad de Valladolid (Spain)

Résumé : Ce travail a un double objectif : expliquer certaines limites ayant une influence sur les approches traditionnelles de la théorie de la réussite, et souligner un critère pour l'évaluation comparative des succès empiriques de la théorie. On insiste sur les points suivants : a) la supériorité de la prédiction sur l'adaptation, b) la résolution des anomalies non-réfutées et c) l'utilisation limitée des hypothèses *ad hoc*. Après une première partie consacrée à la désambiguïsation du label « succès de la théorie », la deuxième section traite de quelques-unes des principales lacunes du traditionnel, une approche conséquentialiste de la théorie de la réussite. Puis quelques éléments sont donnés afin de renforcer le critère de succès empirique de la théorie, souligné dans la dernière section.

Abstract: The purpose of this work is twofold: to explain some of the limitations affecting traditional approaches to theory success, and to outline a criterion for the comparative evaluation of a theory's empirical success. A special emphasis will be placed on the following issues: a) the superiority of prediction over accommodation, b) the resolution of non-refuting anomalies, and c) the limited use of *ad hoc* hypotheses. After a first section devoted to the disambiguation of the label "theory success", the second section discusses some major shortcomings of traditional, consequentialist approaches to theory success. Then some clues are provided so as to strengthen the criterion for a theory's empirical success, which is outlined in the last section.

1 Preliminary remarks: the disambiguation of "theory success"

The starting point of the paper is a distinction between four main, not mutually exclusive kinds of theory success: 1) the one concerning a greater empirical adequacy or fit (oxygen theory versus phlogiston theory, relativistic

Philosophia Scientiæ, 19(1), 2015, 1–14.

mechanics versus Newtonian mechanics), 2) that due to an increase in predictive power, which clearly entails a rise in empirically "promising" informativeness (Relativity Theory [theory of relativity?] versus Newtonian mechanics), 3) that derived from a higher explanatory capacity accomplished in terms of a more detailed specification of those causal mechanisms underlying the empirical phenomena under study (Mendelian genetics versus theory of the mixture), 4) the one related to a higher explanatory capacity achieved through a more systematized and unified account of empirical phenomena (Newtonian mechanics versus Galilean Mechanics). Theory success of either kind can be understood as a form of explanatory success (predictive, informative, causal, and unitary-systematic respectively), each of them having been historically emphasized from different philosophical standpoints (by authors like K. Hempel, P. Kitcher, M. Friedman, and W. Salmon). This persistent ambiguity surrounding the notions of explanation and success will be avoided here by carefully specifying which side of explanation is being considered. As already said, I am going to focus on the first kind of theory success, which, from the approach adopted here, presupposes also the second kind of success.

The above distinction, however, becomes more complicated once that [delete?] we take into account the fact that either kind of theory success may occur in a (conceptually) continuous or discontinuous fashion, depending on whether the successor (more successful) theory is conceptually compatible (or commensurable) with the predecessor (less successful) one. Challenges posed by conceptual discontinuity will be minimized here by assuming T. S. Kuhn's late notion of local incommensurability [Kuhn 1982, 1993], which allows for empirical commensurability and rational comparison between rival theories. The concepts shared or commonly presupposed by incommensurable theories would provide a commensurable ground for theory evaluation.

Another point to clarify has to do with [is?] the possibility of distinguish[ing?] between success due to a theory and success merely accompanying the theory, that is, taking place as the theory develops but not because of the theory. Throughout this paper, the expression "theory success" will refer to that success dependent on the theory under consideration, rather than to that not dependent on such theory but just occurring simultaneously with its development. The example of the caloric theory is often used to show how certain research components are preserved in science. This preservation has been accounted differently depending on whether the adopted approach is realistic or instrumentalist. Authors [like?] S. Psillos, P. Kitcher or J. Worrall, provide a realistic account of the caloric case, claiming that much progress made by the caloric theorists has been preserved by upholders of succeeding theories. According to Psillos, the achievements made during the development of the caloric theory, were the following: the development of calorimetry (specific heat), the law of adiabatic expansion of gases, and Carnot's theory of heat engines [Psillos 1994]. On the other hand, non-realists like L. Laudan or H. Chang reply by objecting that what is preserved does not include theoretical components but elements that are independent of the caloric theory. In emphasizing the independence of these research constituents, Chang stresses the relevance of the following: observational data, phenomenological laws, non-empirical elements like representational or inferential techniques (including mathematical methods), and deeply rooted metaphysical commitments [Chang 2003, 910–911]. This metaphysical debate on preservation will not be resumed here, since the present focus of discussion is rather the dependence or independency of research achievements with respect to a given theory. Whatever view we hold on this particular case, the interesting point to note here is the very possibility that not all achievements accomplished during the period when a theory is developed are due to the very theory.

Finally, questions about whether theory success implies some sort of truth approximation, or whether it provides grounds for a realist conception of scientific theories, are going to be put aside. Since arguments in favour of these ideas are supported on considerations concerning success of the above kinds, clarification of the latter turns out important regardless of what view is hold with respect to the former issues. This metaphysical neutrality extends here over all realist and anti-realist options in the "market", from entity realism (different versions of which are endorsed by Kitcher, Hacking, and Giere among others) to structural realism (developed by authors like Worrall, Ladyman, French), and anti-realism in its different varieties and degrees (Laudan, van Fraassen, Cartwright).

2 Limitations of traditional approaches to theory success

As already pointed out, traditional criteria for theory success mainly revolved around the number of a theory's successful applications—which, from the statement view of theories [Popper 1962], amounts to the number of a theory's true empirical consequences, and, from a model-theoretic approach [Moulines, 2000: missing in the reference list], to the number of phenomena successfully embedded into theoretical models through those models' empirical substructures. In characterizing a theory (T_2) as more successful than another theory (T_1), traditional accounts of theory success (S) have usually commit to [commited?] the following, merely quantitative condition: $S(T_1) \subset S(T_2)$, sometimes supplemented with the Kuhnian condition that $S(T_2)$ includes some of T_1's most recalcitrant anomalies.

However, this purely quantitative account of theory success, even if supplemented by [with?] the assumption that there is an inter-theoretical epistemic-pragmatic criterion for prioritising certain intentional applications whose possible extension to models is of special interest, is not revealing enough as an analysis of theory success. A major problem has to do with the fact that the requirements placed on the success of the [delete?] competing theories are merely quantitative, and for this reason, insensitive both to the empirical import of the intended applications covered by the theories and to the explanatory significance of the models accounting for the applications. With respect to the first, it is worth mentioning the special relevance given to novel, unexpected predictions or to those concerning salient phenomena which were not initially included in the domain of application of the theories. It seems plausible to think that even if a theory fails to satisfy the condition mentioned in [the?] last paragraph, it could still be considered as more successful than its competitor in case that only the first provided unexpected predictions and an explanation for previously disregarded salient phenomena. Relativity theory may illustrate this kind of success, since it was considered successful, at least partly for these reasons, even before it could be confirmed to the same extent as Newtonian mechanics had been. There can also be cases (like Ptolemaic and Copernican theories) in which rival theories do not differ in their quantitative success, and yet one of them is considered as explanatorily superior to the other. The two conditions above are not applicable here, and yet there seems to be a clear difference in success between the two theories. Before Galileo, Ptolemaic and Copernican astronomy could roughly account for the same phenomena. However, the second was regarded as a better explanation, not only due to its greater simplicity and unifying capacity, but also to its less *ad hoc* character. A further problem affecting the traditional standpoint concerns the meagre attention paid to the scope, generality or informativeness of a theory as an element of success. In other words, the quantitative requirement can be met without the scope of the theory significantly changing.

The above observations suggest that some important questions remain unanswered in traditional accounts of theory success. These questions point to the need for refinements in the form of several qualitative requirements, which should be specially focused on issues regarding a theory's scope and possible *adhocness* [maybe it would be worth hamonizing the way you write ad-hoc-ness?]. More in particular, such requirements will concern: a) the resolution of non-refuting anomalies [Laudan 2000, 166–167] the superiority of prediction over accommodation [Lipton 1991, 68], and c) the limited use of non-corroborated auxiliary hypotheses [Thagard 1978, 86–89]. Notice the especial relevance that all the above features have to the question of empirical adequacy, and which becomes evident in the fact that a) is directly connected to both empirical adequacy and informativeness; b) is related to informativeness and indirectly to empirical adequacy (possibility of increasing empirical adequacy by increasing informativeness); and c) is indirectly related to the empirical adequacy of a theory, since it is directly related to the empirical adequacy to the theory's auxiliary hypotheses.

In order to provide an answer to these issues, a stronger notion of empirical adequacy is sketched in the next section.

3 Developing a stronger notion of empirical adequacy

In dealing with the problem of theory success, traditional approaches face some serious shortcomings, in particular those related to the neglect of the problems regarding ad hoc-ness or the insufficient attention paid to issues concerning the application openness or incompleteness of a theory. A stronger notion of empirical adequacy is needed in order to overcome these shortcomings. So next we are going to take into account several hints provided by some of the authors who most seriously attempted to show the significance of the above qualitative requirements. Let us first pay attention to the issue concerning non-refuting anomalies.

A) The notion of non-refuting anomaly is introduced by Laudan in *Progress and its Problems* [Laudan 1977], where he characterizes anomalies in general as "empirical problems which raise reasonable doubts about the empirical adequacy of a theory" once another theory has solved them [Laudan 1977, 28, 30]. According to him, non-refuting anomalies, in contrast to refuting ones, do not involve any logical incompatibility between empirical consequences of the theory on the one hand and verified statements regarding empirical facts on the other [Laudan 1977, 27–29]. They rather entail a theory's incapability to account for certain kind[s?] of salient empirical phenomena whose description is consistent with everything established by the theory. As Laudan puts it:

> Such non-refuting anomalies typically arise when one theory is compatible with, but offers no solution to (or explanation of), certain phenomena for which some [of?] its rivals can give an account. It was my claim that the scientific methodologies of logical empiricism had not recognized this historically significant form of anomaly. [Laudan 1981, 618]

Non-refuting anomalies, therefore, do not primarily point to any mistake on the way in which a theory explains the phenomena but rather to the incompleteness on the part of the theory. Typical cases in which this happens are the ones pointed out by Laudan when developing the notion of non-refuting anomalies in arguing for the importance of completeness as a theoretical virtue. In his work from [of?] 1977 he mentions two cases: the incapability of pre-Galilean kinematics to explain the mathematical features of pendular motion, i.e., the absence of predictions for the geometry of the moving weight, and Newtonian mechanics' lack of explanation for the coplanarity and common direction of the planets' orbits, which had been accommodated in Keplerian and Cartesian astronomies [Laudan 1977, 29]. Some other examples are added in his 2000 paper; for instance, the fact that continents fit together, for which stable-continent theories of geology offered no explanation, or the phenomenon of residual background radiation, which remained unexplained by steady state cosmology [Laudan 1977, 167].

Laudan's conception of anomalies, strongly inspired by Kuhn's [1962/1970, 52-65] [missing in the reference list],[1] includes the idea that anomalies can only be regarded so when another theory has been capable of solving them.[2] Here a wider notion of anomaly is favored instead. According to this wider characterization, anomalies (of either kind) consist in empirical problems that raise rational doubts about the empirical credentials of a theory regardless of whether another theory has succeeded in solving them.

We can already give an idea of what distinguishes non-refuting anomalies from cases where some phenomena remain unexplained by a theory either due to the fact that they fall outside its domain of application or because, despite being included in the domain, it is foreseeable that the theory, as it stands, will be able to account for them in the future. Unlike the first kind of case, non-refuting anomalies have close similarity or connection to paradigmatic exemplars of a theory's intended applications; as opposed to the second kind of case, these anomalies, because of their recalcitrant nature, suggest that the theory will need substantial modifications, additions or even a replacement. Against this, and favoring Laudan's view, it could be argued that unless there is a contrast class of alternative theories capable of solving certain non-refuting empirical problems, we do not have grounds for regarding the latter as anomalies rather than just mere problems of adjustment between theory and data, or, alternatively, mere empirical findings not targeted for explanation (even if closely similar to paradigmatic exemplars). In reply to the first option, it must be emphasized that non-refuting anomalies point not to those cases in which a theory speaks only very approximately but to those cases where a theory, despite possible efforts to the contrary, remains silent with respect to certain phenomena. As for the second possibility, it is important to note that anomalies consist in salient or striking phenomena that, given their relevance to the theory, the latter should be able to accommodate.

Let us now turn to the most central and insightful aspect in such characterization, i.e., the widening of the evidential or evaluative basis for a theory to embrace verified empirical statements that are neither among the set of the theory's consequences nor among the set of the theory's excluded conse-

1. It may be worthwhile to briefly recall Kuhn's notion of anomaly, since some of Laudan's points were already suggested by the former, who nevertheless failed to fully realize about their consequences for the traditional conception of evidential support. The general notion of anomaly introduced by Kuhn corresponds to those problems or phenomena that a theory cannot accommodate and that do not fit the theoretical expectations [1962/1970, 58]. Both refuting and non-refuting anomalies fall under the above general notion. Finally, Kuhn, as opposed to Laudan, does not regard it as necessary for an anomaly to be recognized as such that some rival has been able to solve it. On the contrary, he argues that it is the previous awareness of anomaly what [that?] initiates the process of theory modification or theory change [1962/1970, 62].

2. Cf. [Laudan 1977, 29, also n. 15]. Consequently, Laudan equates what has been called "Kuhn's losses" [Kuhn, 1962/1970, 107-108] with certain instances of non-refuting anomalies, namely, those in which the successor theory provides no explanation for phenomena that the previous theory successfully covered.

quences. To use Laudan's own terms, his discussion of non-refuting anomalies entails the rejection of the "consequentialist theory of evidence or plausibility" [Laudan 1995, 28], as well as the recognition that "[...] being a consequence of a hypothesis is neither necessary nor sufficient to qualify something as evidence for that hypothesis" [Laudan 1995, 29]. The warrant conditions of a statement, therefore, should not be equated with its truth conditions, since poor explanatory power would raise doubts about the epistemic virtue of a theory regardless of whether the latter's truth conditions are widely satisfied [Laudan 1995, 33].[3]

The empirical adequacy of a theory, then, does not only depend on the latter's empirical consequences being true but also on them corresponding to the most salient phenomena in its domain of application. In model-theoretic terms: a theory's empirical adequacy does not only require that (in at least one of its models) all its empirical substructures are isomorphic to the corresponding phenomena [van Fraassen 1976], but also that they are so to all salient phenomena in the domain. The challenge, then, is to characterize the kind of information that, even if logically disconnected from what a theory entails, nonetheless provides crucial evidence for the theory and falls inside its domain of application. From the examples chosen by Laudan, it seems that he has in mind cases in which some striking empirical regularities remain theoretically unexplained, even though they clearly fall within the theory's domain. As we will see in section 4, the threefold evaluation of theories put forward by Kuipers involves the consideration of what he calls "neutral results" [Kuipers 2005], which amount to Laudan's non-refuting anomalies and represent an important parameter in the comparative assessment of theory success.

B) Let us now turn to the issue concerning the priority of prediction over accommodation [Lipton 1991, 68], [Forster 2000, 233]. Intuitively, it seems clear that the quality of the evidence provided by predictions is better than that gathered through accommodation. At first glance it appears that this difference in the quality of evidence lays in the way a predictive theory is devised and tested in contrast to how an accommodative theory would be built and evaluated. In the first case, an *ad hoc* conception of the theory would be more difficult, since the evaluation of the theory would depend on what the theory establishes either about future events or about past events

3. In his 1988 paper, T. Nickles argues that the consequentialist model of scientific justification should be combined with Laudan's generative model, since the second point to theoretical changes that fall outside the standard conditionalization, which would depend on background knowledge remaining fixed [Nickles 1988, 10]. Although in a different context such as the field of mathematics, I. Lakatos introduced a notion similar to Laudan's non-refuting anomalies, namely, that of *heuristic* falsifiers. He explains that, unlike logical falsifiers, which show that a theory as such is false (inconsistent), heuristic falsifiers merely show that a theory does not *explain* properly what it set out to explain—it is a false theory of the informal domain in question. Still, when Lakatos claimed that "*the crucial role of heuristic refutations is to shift problems to more important ones,* to stimulate the development of theoretical frameworks with more content" [Lakatos, 1967/1978, 40], [Lakatos 1978].

that, even if already known, were not initially targeted by the theory. An accommodative theory, by contrast, could be built just to fit events already included in its initially targeted domain.

The above account of the difference between prediction and accommodation, however, is neither completely right nor very enlightening. It is not completely right because it presents accommodation as merging into *ad-hoc*-ness, when the former does not necessarily imply the latter. The Darwinian theory of evolution, for instance, despite its [it?] being more accommodative than predictive, is not regarded as an *ad hoc* theory. The intuitive account of accommodation is not very enlightening neither, since, regardless of how a theory was initially devised (in an *ad hoc*, accommodative or predictive manner) it may turn out something different in the future depending on different dynamical aspects concerning theory testing as well as theory evolution. As argued by Peter Lipton, lack of *ad-hoc*-ness and precision should be emphasized as the most distinctive features of predictions as opposed to accommodations. The implications of *ad hoc*-ness will be discussed in the following section. As for precision, one of the main reasons why predictions are precise is that, as opposed to accommodations, they can be subject to experimental control [Lipton 1991, 169]. Anticipating unknown facts, on the other hand, even if clearly constituting an advantage of predictions, does not represent one of its most essential features. Not surprisingly the use of Einstein's special relativity theory to predict the deviation of Mercury's perihelion is also mentioned as an instance of a prediction providing stronger support than mere accommodation, despite the fact that such deviation was an already known phenomenon. The soundness of background knowledge and auxiliary assumptions is also mentioned as an indicator of non-*ad-hoc*-ness.

Put in a nutshell, as opposed to what happens with predictive theories, in the case of accommodative theories, the domain of phenomena that prompted the construction of a theory is not different from the domain of phenomena providing the evidential basis to test the theory. This suggests the success condition that the evidential domain of a theory be different from its "construction" domain. This, in turn, would entail a higher independece of the evidence with respect the theory for which it plays the evidential role, for the evidence involved in testing the theory would not be part of the domain of phenomena that the theory meant to explain.

C) Let us turn now to another condition for empirical adequacy such as it is the limited use of *ad hoc* hypotheses, i.e., non-corroborated auxiliary hypotheses. P. Thagard has established this condition as one of the simplicity criteria for theory choice [Thagard 1978, 86–89], yet in so far as the condition essentially concerns corroboration it can also be understood as a condition to further qualify the empirical adequacy of a theory. According to this author:

> An *auxiliary hypothesis* is a statement, not part of the original theory, which is assumed in order to help explain one element of

> F [the facts] or a small fraction of the elements of F. [Thagard 1978, 86, italics in the original]

Before this characterization, Thagard has already excluded the possibility that A (i.e., a set of auxiliary hypotheses) be equated with C, that is, a set of unproblematic conditions accepted independently of T or F and assumed in T's application. Instead, A would include assumptions with narrow applications, restricted in use to one class of fact. As examples, he mentions Huygens' assumption, in order to explain the irregular refraction in Iceland crystal, that some waves are spheroidal rather than spherical, or his assumption that the speed of light is slower in denser media to explain Snell's law of refraction. Thagard points out that these assumptions sometimes coincide with *ad hoc* hypotheses, although he notes that, since *ad hoc*-ness is a dynamic property, auxiliary assumptions may lose this *ad hoc* status at some point. This may occur either because they are at some point confirmed or due to the fact that competing theories share the same assumptions.

After severely scrutinizing most common accounts of *ad hoc*-ness, both J. Leplin and M. Carrier have separately argued that none of such accounts sufficiently emphasizes the truly distinctive feature of *ad hoc* hypotheses. Both agree that a hypothesis' *ad hoc* character is not determined by the way it was devised but rather by the fact that it does not receive independent empirical support. The two [both?] authors converge in pointing to excess empirical content as a key requirement for non-*ad-hoc*-ness. In Carrier's terms:

> A hypothesis explains a fact in a non-ad-hoc manner, if it simultaneously explains at least one additional independent fact that either constitutes an anomaly for the rival theory or that falls beyond its realm of application, i.e., that is neither derivable from nor inconsistent with the competing approach. [Carrier 1988, 206]

Although implying some further requirements, J. Leplin's definition of *ad-hoc*-ness also includes a condition concerning the lack of independent empirical support [Leplin 1975, 336–337]. Yet, his most novel contribution to the analysis of *ad hoc*-ness consists in a condition regarding non-fundamentality. This condition reveals some important aspects related to *ad hoc*-ness—like the locally holistic nature of anomalies and the corresponding requirement for hypotheses to solve several anomalies together. The condition, however, also entails some questionable points, specially the presupposition that non-fundamental theories, i.e., those with a wide variety of serious insufficiencies, are the only ones affected by recalcitrant *ad-hoc*-ness. Given the dynamical nature of theory justification,—which, as recognized by both Leplin and Carrier makes *ad hoc*ness a dynamical, non-stable property—, it may be equally possible that a non-fundamental theory eventually overcomes its difficulties and becomes a fundamental one. Thus, Leplin's distinction between incomplete theories (where *ad hoc* hypotheses are eventually changed into non-*ad-hoc* ones) and non-fundamental theories (where the above possibility is ruled out) does not turn out very helpful to elucidate the question of *ad hoc*-ness. On the one hand, we can never be sure that an incomplete theory is not going to end up revealing itself as non-fundamental and, conversely, a non-fundamental one revealing as the opposite. That is the reason why, most often, unanimous judgments about whether a theory is of one kind or the other are made only when a theory has been actually completed or replaced, and yet unanimous judgements about *ad-hoc*-ness can be made before that happens. Given its questionable character, non-fundamentality is not included here as a necessary condition for *ad-hoc*-ness. In other words, a general criterion of [for?] *ad hoc*-ness should enable us to compare two theories in terms of their relative *ad hoc*-ness regardless of whether we regard either theory incomplete or non-fundamental.

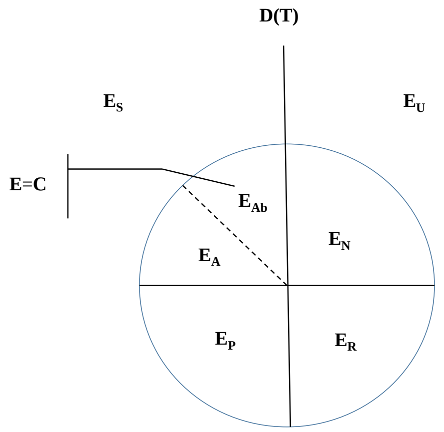
4 Outline of an empirical criterion for theory success

The three qualitative requirements for theory success commented above have all something in common: they imply some conditions concerning the relation between a theory and different subsets of its domain. So, before outlining an empirical criterion for theory success that includes the above requirements, several distinctions between different sub-domains of a theory should be drawn. There are two basic divisions: one between successful and unsuccessful intended applications and the other between the construction domain and the evaluative domain. The latter would include four sub-domains: two respectively corresponding to refuting and non-refuting anomalies, on the one hand, and two corresponding to successful prediction[s?] and accommodations, on the other. For the sake of simplicity, let us introduce the following notational conventions, all of them referring to domain specifications relative to a theory.

D =	domain of application		
C =	construction domain		
E =	evaluative domain		
ES =	successful intended applications	EU =	unsuccessful intended applications
EP =	successful predictions	ER =	refuting anomalies
EA =	successful accomodations	EN =	non-refuting anomalies
EAb =	*ad hoc* accomodations.		

The corresponding sub-domains of a theory are represented in the diagram below. As indicated in the [missing word?] there, only in the case of *ad hoc* hypotheses the construction domain and the evaluative domain of the theory completely overlap.

The empirical criterion for theory success can now be outlined as follows:



T2 is more empirically successful than T_1 if the following conditions hold:

a) $E_S(T1) \subseteq E_S(T_2)$,

b) $E_N(T_1) \cap E_S(T_2) \neq \varnothing$,

c) $E_A(T_1) \cap E_P(T_2) \neq \varnothing$,

d) $||E_{Ad}(T_2)|| < ||E_{Ab}(T_1)||$.

Note that these four conditions are here presented as globally sufficient for comparative empirical success, but not as globally necessary, which would reveal as too strong a requirement for most cases of theory choice. As shown in the diagram, in cases of *ad hoc* accommodations the construction domain and the evalutative domain completely overlap. Non-refuting anomalies of a theory, on the othe hand, are always included within the sub-domain formed by its unsuccessful intended applications.

This approach to theory success can be seen as supplementing Kuiper's complex evaluation matrix by adding some qualitative requirements which are absent from his proposal. In his symmetric models of separate hypothetico-deductive method[s?] of evaluation, i.e., the micro- and the macro-models, for a theory to be at least as successful as the old one, some general conditions of adequacy must be satisfied not only for the definitions of "success" and "problem", but also for that of "neutral result"—which equates to the notion of non-refuting anomaly. Contrary to what happened in the asymmetric models, where success conditions only refer explicitly to individual problems and general successes while neutral results remain hidden, in the symmetric model the latter play an important role [Kuipers 2005, 52]. By taking all three types of results explicitly into account, Kuiper's symmetric models meet Laudan's evaluative requirements concerning the wider scope of most successful theories. Keeping in mind that neutral general facts for a theory constitute neither a problem nor a success, a successful theory should transform general problems into neutral facts (or even successes) and neutral general facts into successes. The same kind of requirement[s?] should be met at the micro-level for individual successes, individual problems and neutral results. There also neutral instances may remain neutral or become positive.

As pointed out earlier, the fine-grained analysis developed by Kuipers fails to incorporate some qualitative parameters—like *ad hoc*-ness and accommodation—which prove essential in the evaluation of theory success.

5 Concluding remarks

Traditional accounts of theory success have proven in a sense too broad and in another too narrow. The first point becomes evident in the neglect of the problems regarding *ad hoc*-ness and accommodation, the second in the insufficient attention paid to issues concerning the application openness or incompleteness of a theory. All these questions, however, seem crucial for a comparative appraisal of theory success. Problems concerning accommodation and *ad hoc*-ness would require an analysis of a theory's level of precision and excess empirical content with regard to its construction domain. Issues regarding narrowness of scope would call for the recognition of non-refuting anomalies as part of the evaluative domain of a theory. The above factors should be taken into account if a theory's empirical success is to be evaluated in all its complexity.

Acknowledgments

I am thankful to Thomas Mormann for helpful comments and references on the subject of theory success. Thanks also to Otávio Bueno for illuminating discussion on the problem of non-refuting anomalies. This work was financially supported by the Spanish Ministry of Economy and Competitiveness through the project *Pragmatics as the driving force behind the study of semantic flexibility: conversational contexts and theoretical contexts* (FFI2012-33881).

.

Bibliography

Carrier, Martin [1988], On novel facts. A discussion of criteria for non-ad-hoc-ness in the methodology of scientific research programmes, *Zeitschrift für allgemeine Wissenschaftstheorie*, 19(2), 205–231, doi:10.1007/BF01801335.

CHANG, Hasok S. [2003], Preservative realism and its discontents: Revisiting caloric, *Philosophy of Science*, 70(5), 909–912.

FORSTER, Malcolm R. [2000], Hard problems in the philosophy of science: Idealisation and commensurability, in: *After Popper, Kuhn and Feyerabend*, edited by R. Nola & H. Sankey, Dordrecht: Springer, *Australasian Studies in History and Philosophy of Science*, vol. 15, 231–250, doi:10.1007/978-94-011-3935-9_9.

KUHN, Thomas N. [1982], Commensurability, comparability, communicability, in: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 2, 669–688.

—— [1993], Afterwords, in: *World Changes: Thomas Kuhn and the Nature of Science*, edited by P. Horwich, Cambridge, MA: MIT Press, 311–341.

KUIPERS, Theo A. F. [2005], The threefold evaluation of theories. A synopsis of *From Instrumentalism To Constructive Realism. On Some Relations Between Confirmation, Empirical Progress, And Truth Approximation*, in: *Confirmation, Empirical Progress, and Truth Approximation*, edited by R. Festa, A. Aliseda, & J. Peijnenburg, Amsterdam; New York: Rodopi, *Poznan Studies in the Philosophy of the Sciences and the Humanities*, vol. 83, 23–85.

LAKATOS, Imre [1978], *Mathematics, Science and Epistemology*, vol. II: Philosophical Papers, Cambridge: Cambridge University Press, edited by J. Worrall and G. Currie.

LAUDAN, Larry [1977], *Progress and Its Problems*, Berkeley: University of California Press.

—— [1981], Anomalous anomalies, *Philosophy of Science*, 48(4), 618–619.

—— [1995], Damn the consequences!, in: *Proceedings and Addresses of the APA*, vol. 69, 27–34.

—— [2000], Is epistemology adequate to the task of rational theory evaluation?, in: *After Popper, Kuhn and Feyerabend*, edited by R. Nola & H. Sankey, Springer Netherlands, *Australasian Studies in History and Philosophy of Science*, vol. 15, 165–175, doi:10.1007/978-94-011-3935-9_6.

LEPLIN, Jarrett [1975], The concept of an *ad hoc* hypothesis, *Studies in History and Philosophy of Science – Part A*, 5(4), 309–345, doi:10.1016/0039-3681(75)90006-0.

LIPTON, Peter [1991], *Inference to the Best Explanation*, New York: Routledge, 2004.

NICKLES, Thomas [1988], Truth or consequences? Generative versus consequential justification in science, in: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1988, 393–405.

POPPER, Karl [1962], Truth, rationality and the growth of scientific knowledge, in: *Conjectures and Refutations*, New York: Harper & Row, 215–252.

PSILLOS, Stathis [1994], A philosophical study of the transition from the caloric theory of heat to thermodynamics: Resisting the pessimistic meta-induction, *Studies in History and Philosophy of Science – Part A*, 25(2), 159–190, doi:10.1016/0039-3681(94)90026-4.

THAGARD, Paul R. [1978], The best explanation: Criterion for theory choice, *The Journal of Philosophy*, 75(2), 76–92.

VAN FRAASSEN, Baas C. [1976], To save the phenomena, in: *The Philosophy of Science*, edited by R. Boyd, P. Gasper, & J. D. Trout, Cambridge, MA: MIT Press, 187–194, 1992.