Idealization, Abduction, and Progressive Scientific Change[1]

Xavier DE DONATO RODRÍGUEZ

BIBLID [0495-4548 (2007) 22: 60; pp. 331-338]

ABSTRACT: After a brief comparison of Aliseda's account with different approaches to abductive reasoning, I relate abduction, as studied by Aliseda, to idealization, a notion which also occupies a very important role in scientific change, as well as to different ways of dealing with the growth of scientific knowledge understood as a particular kind of non-monotonic process. A particularly interesting kind of abductive reasoning could be that of finding an appropriate concretization case for a theory, originally revealed as extraordinarily successful but later discovered to be strictly false or only approximately or ideally true. I try to show this with the example of the Kepler-Newton relation. At the end of the paper, I give criteria in order to construe abductive explanations in correspondence with a reasonable account of empirical progress.

Keywords: abduction, non-monotonic reasoning, idealization, approximation, Kepler-Newton case, empirical progress, scientific change.

Aliseda's book presents many new and enriching perspectives on abductive reasoning having to do not only with the logical structure of abductive inferences, but also with the relation of abduction to many important concepts in philosophy of science, such as scientific explanation, scientific change and scientific progress. It provides not only a logical study which tries to characterize the process of abduction, but establishes a basis for its application to the study of many other concepts, making the work an interesting reading for people from many different areas (philosophy, logic, psychology, linguistics, mathematics, artificial intelligence, or computer science). Among the specificities that I want to mention are the use of the structural logical approach, started by Scott (1971) and applied later by Gabbay (1985) to the study of non-monotonic reasoning and by van Benthem (1996) to logical dynamics, which aims to characterize the basic combinatorial properties of a certain process by using logical structural rules. This tool is used by Aliseda in a satisfactory way not only in the analysis of abduction, but also in the analysis of the well-known Hempelian models of scientific explanation (D-N and I-S models), obtaining the interesting consequence according to which the Hempelian model can be viewed as a particular form of abductive reasoning. This is her main concern in Chapter 5. Aliseda shows the limits of the purely structural rule analysis she uses in Chapter 3 to characterize abductive inferences. Restricted forms of monotonicity and cut rules are given as well as different interesting logical rules. According to Aliseda (2006, p. 137): "It seems clear that the Hempelian deductive model of explanation is closely related to our proposal of (abductive) explanatory argument, in that it complies with most of the logical conditions discussed in chapter 3."[2] Aliseda applies also her analysis to a model of progressive scientific explanations, proposed by

[1] I thank Atocha Aliseda and Larry Laudan for their comments on an earlier draft of this paper.

[2] Aliseda's reading of Hempel's account is that the consistency requirement is presupposed.

Kuipers (1999) and (2000), in order to generate new theoretical explanations for "non-refuting anomalies" —a term coined by Laudan (1977)— (that is, cases which neither confirm nor refute the theory and for which the theory cannot provide an explanation) and more successful theories. The analysis of this notion of scientific progress is in terms of semantic *tableaux*, which had been used until now basically as a refutation method and only applied to the study of abductive reasoning by Mayer and Pirri (1993). The basic idea is that the abductive method can be represented as a process of *tableaux* extension. The author shows that this method, which deploys the idea of computational generation of abductions based on algorithmic procedures, can equally be applied to the theory of epistemic change and belief revision, basically proposed by Gärdenfors, with the result that abduction appears to be a particular kind of belief revision.[3] As far as I know, although the role of belief revision in abduction had already been the theme of Jackson (1990), Aliseda's original contribution is to provide an abductive model for belief revision in terms of semantic *tableaux*.

Apart from the use of formal tools, there are two basic ideas in Aliseda's work, which I find very stimulating: (i) the idea that abduction is more a context-dependent practice which comprises different forms of logical inference and computational processes rather than a new notion of inference, and (ii) the idea that abduction is a method for generating hypotheses in scientific inquiry as well as in the ordinary life (it does not have to do only with the analysis of scientific abductive method, but also with the abductive process implicit in many forms of ordinary reasoning).

But in the present contribution, I want to relate the concept studied by Aliseda to another concept which also occupies a very important role in scientific change: idealization (and its converse: concretization). The author, in the first chapter of her book, recovers an old dispute around the question *does scientific discovery have a logic?* Traditional answers to this question were dogmatic to some extent. Aliseda's answer is not: "there is no precise methodology by which scientific discovery is achieved, as a form of problem solving, it can be pursued via several methodologies." (Aliseda 2006, p. 16) —abduction being one of them, of course. I agree with her in this point, but I want to be provocative and propose the concept of *idealization* as a key to understand the process of scientific inquiry and discovery, in fact as "the very essence" of scientific method —if we can still put the thing in these terms.

Idealization is here understood not only as model construction but also as the assumption of certain counterfactual conditions in order to derive laws, from a theory which is considered to be more correct, which are thought only approximately correct or ideally true. In cases of approximative reduction of a theory from another one or in cases of derivation of laws and theories according to certain ideal conditions, there occurs a process of progressive theoretical change in which the replacing theory should explain the successful empirical applications of the old theory and, at the same time, explain where it failed. In terms of Glymour (1970, p. 341), these are cases in which "a theory is explained by showing under what conditions it would be true, and

[3] See the pioneering work of Alchourrón, Gärdenfors and Makinson (1985) as well as Rott (1991) and Gärdenfors and Rott (1995).

by contrasting those conditions with the conditions which actually obtain." This notion differs, at least in principle, from the classical notion of deductive intertheoretical explanation, which is the core of a well-known conception of scientific growth —let me call it "the naïve conception." According to this, science grows in a cumulative way, following a linear order, scientific evolution being *monotonic*. According to the approach sketched above –in terms of idealization and approximation—, the growth of science is essentially *non-monotonic*. One should probably say that science can grow as a consequence of *both* kinds of evolution. The incorporation of the empirical laws of magnetism and electricity (Coulomb, Gauss, Biot-Savart, Ampère, and Faraday) into a concise mathematical form made by Maxwell about 1860 is an example of monotonic growth, though it is also true that Maxwell had to "correct" Ampère's law by introducing the notion of the electrical displacement current in order to make it more general: valid not only for the case of a continuous current, but also for cases of discontinuous current. In fact, he did not really correct it: he merely made the law more general or, better yet, he formulated a more general version of the law. The original Ampère's law remains valid for a restricted domain, which is not the case of the Kepler's laws, which are strictly false. The Kepler-Newton case (Kepler's laws of planetary motion vs. Newtonian celestial mechanics) or the Newton-Einstein case (classical celestial mechanics vs. theory of relativity) would be examples of non-monotonic change, in which idealization and approximation play an essential role.

It is the logic of idealization which Rott (1990) and (1991) tries to capture in his analysis. To do it, he bases his analysis on a model developed by Peter Gärdenfors to account for the counterfactual suppositions which determine the change of belief and the "revision of theories". According to the classical model of explanation, a theory explains something if the *explanandum* is entailed by the theory together with some additional or applicability conditions. Rott considers three possibilities for these applicability conditions: applicability conditions *C* made by the supporters of a theory *T*:

(1) may be known to be true, realistic (from the standpoint of *T*), or
(2) may be known to be false, idealizing (from the standpoint of *T*), or
(3) it is not known in *T* whether they are true (or they are true or false depending on certain further conditions).

These distinctions yield three conceivable ways to express that a theory *T'* is a *superior successor* for theory *T*:

(RE = real explanation): *T'* provides us with a *real explanation* of *T*: "because *C* is true, *T* is (strictly) correct; but if *C* were not true, *T* would not be (strictly) correct".

(IE = idealizing explanation): *T'* provides us with an *idealizing explanation* of *T*: "if *C* were true, *T* would be correct; but because *C* is not true, *T* is not correct".

(CE = conditional explanation): *T'* provides us with an *conditional explanation* of *T*: "if *C* is true, *T* will be correct; but if *C* is not true, *T* will not be correct".

According to Rott (1990, p. 110-111):

T’ is a *superior successor* theory *T* iff there are conditions *C* for *T* such that the minimal revision of *T’* needed to incorporate $\begin{Bmatrix} C \\ T \\ non-C \\ non-T \end{Bmatrix}$ into *T’* includes $\begin{Bmatrix} T \\ C \\ non-T \\ non-C \end{Bmatrix}$.[4]

Rott imposes a strict condition on the notion of minimal revision: *T’* cannot be a superior successor of *T* if neither *C* nor *T* nor non-*C* nor non-*T* *overthrows the fundamental axioms of T’*. The Kepler-Newton case, for example, is an instance of idealizing explanation in the sense just explicated, Newton’s theory being a superior successor. According to Rott’s analysis, the applicability conditions cannot include the ideal supposition that the masses of the planets are zero (a common idealization present in many books), just because it is irreconcilable with Newton’s theory. The reason: the strict condition imposed on the notion of minimal revision.[5]

The Kepler-Newton case and other equally important examples of intertheoretical relations can be also reconstructed as approximations of certain kind. When we take Rott’s analysis as a complementary view to that of approximation, it seems to me that we’ve got an interesting and adequate way of putting the things. His analysis should not be taken as invalidating the approximative analyses. In fact Rott’s does not succeed in eliminating the approximative aspect of the Kepler-Newton relation: one of his counterfactual assumptions involves an approximation with regard to the mass of the planets (the common idealization according to which the mass of the planets is zero in comparison with that of the Sun implies in fact an approximation relation). So, my first critical point with regard to his analysis is that he does not succeeded in avoiding the approximative aspect of the intertheoretical relation. This is a problem because there are also good reasons to reconstruct this and other similar cases as approximation relations. Intertheoretical approximation has received a rigorous formal treatment in Balzer, Moulines and Sneed (1987) and Pearce and Rantala (1984), both having reconstructed the Kepler-Newton case, the first in the structuralist terms of the Sneed-Stegmüller approach and the second in terms of model theory and non-standard analysis. My own approach tries to relate both the approximative and the idealizational component in a new formal characterization in model-theoretic terms (see de Donato (2005) for details). A second point is that he does not complete his analysis showing the further concretization level, when Newton introduces additional transformations for systems of three or more bodies, such as the Sun-Earth-Moon or the Sun-Jupiter-Saturn systems. At the end of sect. XI of the *Principia* (see Propositions LXV-LXVIII), he shows why the first two laws do not (strictly) hold for multi-body systems: the reason is briefly that there is no single fixed force-center for multi-body systems and, therefore, planetary orbits will approach nearer to an ellipse and the

[4] Rott attempts to show this as a consequence (theorem) of his analysis. See also Rott 1991, p. 254.

[5] See de Donato (2005, pp. 163-170) for a historical and philosophical discussion of the Kepler-Newton case. For Rott’s analysis of this case, see Rott (1990, pp. 113 and ff.) and Rott (1991, pp. 284 and ff). See also Cohen (1980) for an authoritative historical account.

descriptions of the areas will be only approximately uniform if the center be taken as the center of gravity of the system consisting of the Sun and all the planets which are between the Sun and the planet whose orbit is under consideration. Newton considers in this point different cases and employs several approximations.

Kepler also made idealizations to simplify his own laws, in particular he suggested an approximation method, that of the simple equant centered on an empty focus, to simplify the application of his law of areas. This approximative or idealized version was indeed a good approximation, though it was less accurate in the region of the octants. Kepler himself rejected this idealized version of his second law before he had arrived at the genuine focal ellipse.[6] But Kepler was evidently unaware of the idealizations necessary to "derive" his laws from the theory of universal gravitation.[7] As Popper recognizes (1972, p. 201), from the theories of Kepler or Galileo we do not receive any hint of how we can correct them in order to obtain a more general and accurate theory. Only when we have available Newton's theory shall we be able to find out whether the old theories are approximately true, to what degree of accuracy and how can then be obtained from the new theory (which idealizations we need). Nevertheless, when Newton tried to construct his own theoretical hypotheses, that is, when he was searching for the law of universal gravitation, he proceeded according to the *desideratum* of explaining Kepler's Laws of Planetary Motion as a special, approximative case of his own theory. In the same way, Einstein was in search of an adequate formulation of the gravitatory field equations in a way such that he could derive Newton's law of universal gravitation as an approximative case of his theory. When we have the concretised form of a law or theory it is relatively easy to obtain an old and only approximately correct set of laws, but the most important challenge is how to obtain new and more precise laws or theories. Here is where abductive and belief revision processes take place.

Science seems to grow by trying to make more precise the laws and theories we have. Sometimes it is necessary to correct them and empirical investigation leads us to extend the domain of our theories. New theories can be applied to a greater domain of entities and old theories can be shown to be valid for a restricted subclass of this domain when we assume certain ideal conditions. Under these ideal conditions it is then possible to derive old theories from the new ones. A paradigmatic case of this seems to be that of Kepler-Newton. Science grows then by concretization. Concretization involves tentative hypotheses stating whether a certain (new) factor has or does not have an appreciable influence on our computations, on our laws. These concretizations can be made years or centuries after a theory has been formulated, as in the case of Newton trying to derive Kepler's laws from his own gravitational theory. But they can also be made immediately as a way of testing our hypotheses. The idealization-concretization process, understood as the process which leads from idealizations

[6] See Dreyer (1953, p. 390) and I.B. Cohen (1980, p. 225).

[7] As an introduction to the study of these two cases, it seems to me interesting to read I.B. Cohen's (1974) paper for a comparison between the philosophical and the historical points of view in relation to the analysis of scientific theories.

by constructing models to the concretizations which make our theories more accurate, seems to be the essence of scientific method. As this method of making our theories more accurate is inseparable from empirical testing, it then involves observation and experiment as necessary components. The accuracy of our theoretical models must be understood as the degree to which our theoretical models approach the data. Idealizations can then be justified by showing how they lead to more accurate theories by the converse process of concretization: "If it can be shown that more realistic initial conditions will lead via theory to correspondingly more accurate predictions, then the original highly idealized initial conditions are justified in the sense that they provided the starting point for a successful confirmational process" (Laymon 1982, p. 115).[8] The well-known puzzle of how to confirm idealizations can be solved in its turn by invoking the concept of approximation: it is argued that it is "impossible" to test idealizations, because their antecedents (ideal conditions) are never realisable. To this objection we can reply: maybe idealizations themselves cannot be directly confirmed, but the statements in which it is stated that the idealized factors *do not assume the limit value*, but *approach to it*, these statements can usually be confirmed without problem. We merely need experimentally reproduce a situation which approaches that of the ideal or limit conditions and try to show whether the idealized law approximately holds. There is another kind of justification for the use of idealizations: they should allow for practical computability. In intertheorical cases, such as the Kepler-Newton and the Newton-Einstein relations, the practical computability allowed by the idealized theory enable us to use the old laws *in some restricted domains* and *to a certain degree of accuracy*.[9]

To conclude, let me return to the theme of abduction. To abduce the best explanation becomes in the present context to find the best concretised form of a certain law, which is proved to be only approximately true. In many paradigmatic cases of "revolution" in science (for instance Newton and Einstein), we have a new theory which ideally tries (1) to explain the anomalies of the old theory, (2) to explain the same successful empirical consequences as the old theory, (3) to have more predictive power, and (4) if it is possible, to be applicable to a much broader empirical domain of entities. Theories which were well accepted in the past and that are proved to be approximately true according to the new and more precise ones should not be rejected. They can still be applicable in a restricted domain. In these cases, it is desirable that the new theories can provide a theoretical explanation of the fact that the old theory (or certain laws belonging to it) are still successful to some extent. Newton accepted Kepler's laws as empirical generalizations holding only approximately and arrived at his theory of gravitation in part trying to give a theoretical explanation for these laws. Einstein tried to give a formulation for the gravitation field equations which could explain the successful part of Newton's theory. Inference to the Best Explanation

[8] Compare Niiniluoto (1999, pp. 141-143).

[9] A well-known approach to idealization is Nowak (1980). See de Donato (2005) for a critical discussion of the literature on idealization and for a new formal, model-theoretic treatment, which tries to capture the idealizational component in intertheoretical approximations.

(IBE) and abductive processes are normally related with a surprising phenomenon or a dramatic anomaly, both motivating the search for a new explanation in science. The surprising phenomenon should be identified, in this context, with the fact that the old theory cannot explain a certain phenomenon considered highly important by the scientific community. The anomaly is, in our case, the well-observed fact that contradicts the old theory (for instance, the perihelion of Mercury) but, at the same time, is explained by the new one. The problem is then to search for a concretization which ideally satisfies requirements (1)-(4).[10] Maybe there is no specific logic of discovery, but (1)-(4) should be seen as important *desiderata* and as a guide in the search for new and more successful theoretical explanations. I see no ground why these considerations could not be easily taken into account in terms of Aliseda's semantic *tableaux* and of her structural account. To put the things in these terms would be another sign of the richness of her original approach.

REFERENCES

Aliseda, A. (2006). *Abductive Reasoning: Logical Investigations Into Discovery and Explanation*. Berlin: Springer.

Alchourrón, C., P. Gärdenfors, and D. Makinson (1985). "On the logic of theory change: Partial meet contraction and revision functions", *Journal of Symbolic Logic* 50, 510-530.

Balzer, W., C.U. Moulines and J.D. Sneed (1987). *An Architectonic for Science*. Dordrecht: Reidel.

Benthem, J. van (1996). *Exploring Logical Dynamics*. CSLI Publications, Stanford University.

Cohen, I.B. (1974). "Newton's theory vs. Kepler's theory and Galileo's theory: an example of a difference between a philosophical and a historical analysis of science", in Y. Elkana (ed.), *The interaction between science and philosophy*. Atlantic Highlands, N.J.: Humanities Press, pp. 299-338.

——— (1980). *The Newtonian Revolution*. Cambridge: Cambridge University Press.

De Donato, X. (2005). *Idealization and the Growth of Physics*, Doctoral Dissertation, University of Munich, Munich.

Dreyer, J.L.E. (1953). *A History of Astronomy from Thales to Kepler*, Second Edition Revised with a Foreword by W.H. Stahl. New York: Dover Publications.

Gabbay, D.M. (1985). "Theoretical Foundations for non-monotonic Reasoning in Expert Systems", in K. Apt (ed.), *Logics and Models of Concurrent Systems*. Berlin: Springer, pp. 439-457.

Gärdenfors, P., and H. Rott (1995). "Belief revision", in D.M. Gabbay, C.J. Hogger and J.A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 4, Clarendon Press, Oxford.

Glymour, C. (1970). "On some patterns of reduction", *Philosophy of Science* 77, 340-353.

Jackson, P. (1990). "Abduction and counterfactuals", in *Working Notes: AAAI Spring Symposium on Automated Abduction*, Stanford University, pp. 77-81.

Kuipers, T. (1999). "Abduction Aiming at Empirical Progress or Even Truth. Approximation Leading to a Challenge for Computational Modelling", *Foundations of Science* 4, 307-323.

——— (2000). *From Instrumentalism to Constructive Realism*. Dordrecht: Kluwer.

Laudan, L. (1977). *Progress and its Problems*. Berkeley: University of California Press.

Laymon, R. (1982). "Scientific Realism and the Hierarchical Counterfactual Path from Data to Theory", in *Proceedings of the 1982 Biennal Meeting of the Philosophy of Science Association*, ed. by P. Asquith and Th. Nickles. East Lansing, Michigan: vol.1, 107-121.

Magnani, L. (2001). *Abduction, Reason, and Science*. New York, Boston, Dordrecht: Kluwer Academic-Plenum Publishers.

[10] I present these ideas in a sketch form. A similar approach can be found in Kuipers (1999) and (2000). Kuipers studies the role of abduction related to the aim of science as empirical progress, developing a model for instrumentalist abduction and trying to give possible points of departure for computational treatment of this instrumentalist task.

Mayer, M.C., and F. Pirri (1993). "First Order Abduction via tableau and sequent calculi", *Bulletin of the IGPL*, Vol. 1, 99-117.
Niiniluoto, I. (1999). *Critical Scientific Realism*. Oxford: Oxford University Press.
Nowak, L. (1980). *The Structure of Idealization*. Dordrecht: Reidel.
Pearce, D., and V. Rantala (1984). "Limiting case Correspondence between Physical Theories", in W. Balzer, D. Pearce and H.-J. Schmidt (eds.), *Reduction in Science*. Dordrecht: Reidel, pp. 153-185.
Popper, K.R. (1972). *Objective Knowledge*. Oxford: Clarendon Press.
Rott, H. (1990). "Approximation versus Idealization: the Kepler-Newton case", in J. Brzezinski, F. Coniglione, T. Kuipers, and L. Nowak (eds.): *Idealization II: Forms and Applications*, *Poznan Studies in the Philosophy of Science and the Humanities*, Vol. 17, Rodopi, Amsterdam, pp. 101-124.
——— (1991). *Reduktion und Revision. Aspekte des nichtmonotonen Theorienwandels*. Frankfurt a.M.: Peter Lang.
Scott, D. (1971). "On Engendering an Illusion of Understanding", *Journal of Philosophy* 68, 787-808.

Xavier DE DONATO RODRÍGUEZ is Profesor invitado at the Universidad Autónoma Metropolitana de Iztapalapa (Mexico City). From Februar 2006 to Februar 2008 was a Postdoctoral Research Fellow at the Instituto de Investigaciones Filosóficas (UNAM, Mexico City) and PhD Research Fellow at the Ludwig-Maximilians-Universität (Munich, Germany), where he got his PhD with the thesis *Idealization and the Growth of Physics: A Contribution to the Methodology of Science* (2005). His research is mainly in philosophy of science, in particular the study of intertheoretical relations and scientific representation. He has published in *Theoria*, *Crítica* and *Diánoia*.

ADDRESS: Departamento de Filosofía, Universidad Autónoma Metropolitana de Iztapalapa, Avda. San Rafael Atlixco 186. Col. Vicentina, 09340 Iztapalapa, México D.F. E-mail: xavier_donato@yahoo.com.