Ilkka Niiniluoto

"Peirce, Abduction and Scientific Realism"

In: Bergman, M., Paavola, S., Pietarinen, A.-V., & Rydenfelt, H. (Eds.) (2010). *Ideas in Action: Proceedings of the* Applying Peirce *Conference* (pp. 252–263). Nordic Studies in Pragmatism 1. Helsinki: Nordic Pragmatism Network.

ISSN-L 1799-3954
ISSN 1799-3954
ISBN 978-952-67497-0-9

Nordic Pragmatism Network,
Helsinki 2010

www.nordprag.org

Peirce, Abduction and Scientific Realism

Ilkka Niiniluoto
University of Helsinki

1. Introduction

Charles S. Peirce's notion of abduction has been applied as a tool within philosophy of science by many scientific realists. This paper considers the idea that abductive inference can be reformulated by taking its conclusion to concern the truthlikeness of a hypothetical theory on the basis of its success in explanation and prediction. The strength of such arguments is measured by the estimated verisimilitude of its conclusion given the premises. This formulation helps to make precise the "ultimate argument for scientific realism": the empirical success of scientific theories would be a miracle unless they are truthlike. This kind of explanation is not available to those pragmatists who define truth in terms of empirical or pragmatic success. Critical scientific realism seems to give the only viable explanation of the success of science.

2. Peirce and critical scientific realism

Scientific realism as a philosophical position has (i) ontological, (ii) semantical, (iii) epistemological, (iv) theoretical, and (v) methodological aspects (see Niiniluoto, 1999a; Psillos, 1999). It holds that (i) at least part of reality is ontologically independent of human mind and culture. It takes (ii) truth to involve a non-epistemic relation between language and reality. It claims that (iii) knowledge about mind-independent (and mind-dependent) reality is possible, and that (iv) the best and deepest part of such knowledge is provided by empirically testable scientific theories. An important aim

of science is (v) to find true and informative theories which postulate non-observable entities to explain observable phenomena. Scientific realism is thereby in opposition to doctrines like idealism, phenomenalism, positivism, instrumentalism, scepticism, relativism, and social constructivism.

Critical scientific realism can be distinguished from naive forms of realism by two additional theses. First, according to *conceptual pluralism*, all inquiry is always relative to some conceptual framework (just as Kant argued in his critical philosophy), but (unlike Kant thought) such frameworks can be changed, revised, and enriched. Secondly, according to the principle of *fallibilism*, all factual human knowledge is uncertain or corrigible. Even the best results of science may be false, but still they may be truthlike or approximately true.

Conceptual pluralism and fallibilism are characteristic features of the pragmatist tradition as well. However, there is a tension between scientific realism and those forms of classical pragmatism and neo-pragmatism which emphasize that ontology and truth are always relative to human practices (see Pihlström, 1996). This tension was visible when in 1905 Peirce himself attacked some forms of pragmatism in the name of "pragmaticism" (CP 5.411-434). Scientific realists combine fallibilism with the correspondence theory of truth, and thereby reject the so called "pragmatist theory of truth" which *defines* the truth of a belief by its success or utility. Historically, this notion of pragmatist truth was based upon a literal reading of formulations like "an idea is 'true' so long as to believe it is profitable to our lives" in William James's *Pragmatism* (1907), but it can be debated whether James intended this as a definition of truth (see Putnam, 1997; Pihlström, 2008).

Peirce can be regarded as an exponent of critical scientific realism, in spite of his ontological inclination to objective idealism. In his account of the method of science, Peirce emphasized that "real things" affect our senses causally, and in the long run the community of investigators will or would reach "the one True conclusion" (CP 5.384). In my reading, Peirce was not a "convergent realist" in the sense that he would define truth as the limit of inquiry, since he realized that such convergence takes place at best with probability one (CP 4.547n; cf. Niiniluoto, 1984, p. 82). Still, he found it useful to try to characterize the method of science by its ability to "approach to the truth". This was in harmony with Peirce's vision of science as a "self-corrective process" (CP 5.575): "the successful sciences" follow the experimental method which is an application of the rule "By their fruits ye shall know them" (CP 5.465).

Later critical scientific realists have argued – *pace* opponents like W.V.O. Quine and Larry Laudan (see Laudan, 1984) – that it indeed makes sense to say that one hypothetical (even false) theory is "closer to the truth" than another theory. By the same token, it is meaningful to state that a sequence of theories "approaches to the truth", even when the final limit is not reached. Since 1974, after Karl Popper's 1960 attempt to define verisimilitude turned out to fail, the notion of similarity between states of affairs has been employed to give a precise definition of *degrees of truthlikeness* of scientific statements (see Niiniluoto, 1987).

In his fallibilist analysis of inference, Peirce argued that science uses, besides deduction, also two ampliative forms of reasoning: induction and abduction. *Abduction* is reasoning from effects to causes, or from observational data to hypothetical explanatory theories:

(1) The surprising fact C is observed;
But if A were true, C would be a matter of course.
Hence, there is reason to suspect that A is true. CP 5.189

According to Peirce, abduction is "the only logical operation which introduces any new idea" (EP 2:106). It frequently supposes "something which it would be impossible for us to observe directly" (CP 2.640). Against Comte's positivism, Peirce urged in the spirit of scientific realism that science should not be restricted to hypotheses "verifiable by direct observation" (EP 2:225).

Many scientific realists have suggested that the strongest reasons for scientific theories are abductive. Sometimes this idea is connected with a probabilistic account of scientific inference: empirically successful theories are probable. But even when in schema (1) A is the best available theoretical explanation of fact C, it need not generally be the case that A is probable given C. If a fallibilist acknowledges that our strongest theories in science are at best truthlike, then estimated "closeness to the truth" or verisimilitude appears to be a more realistic aim for science than probable truth. Therefore, it is interesting to study the idea that abductive inference (1) can be reformulated by taking its conclusion to concern the *truthlikeness* of a hypothetical theory on the basis of its success in explanation and prediction.

This modification of abduction is also relevant to the so called "ultimate argument for scientific realism". As we have seen, for Peirce the success of science as a fallible cognitive enterprise is based on its method. The most

characteristic feature of the scientific method is its ability to bring the scientific community to a final opinion which is close to the truth. So for Peirce there is an important relation between the success of science and approach to the truth. After the 1950s, when scientific realism became a tenable position after the dominance of empiricism and instrumentalism, several philosophers of science (among them Jack Smart, Hilary Putnam, Grower Maxwell, and Richard Boyd) have defended realism as the best hypothesis which explains the practical (empirical and pragmatic) success of science. The ability of scientific theories to explain surprising phenomena and to yield correct empirical predictions and effective rules of action would be a "cosmic coincidence" or a "miracle" unless they refer to real things and are at least approximately true or truthlike (see Psillos, 1999). It is clear that the form of this "no miracle argument for scientific realism" is abductive (see Niiniluoto, 1984, p. 51).

3. The justification of abduction

Peirce insisted that abduction or "inference to an explanation" has a significant role in science. Often this role has been interpreted as the heuristic function of the *discovery* of new theories, or alternatively as the motive for suggesting or *pursuing* testworthy hypotheses (N. R. Hanson). This is in line with Peirce's methodological characterization of abduction as an "inferential step" which is "the first starting of a hypothesis and the entertaining of it, whether as a simple interrogation or with any degree of confidence" (CP 6.525). The conclusion of (1) states that "there is reason to suspect that A is true". Thus, abduction "only infers a *may-be*" from an actual fact (CP 8.238).

On the other hand, Peirce himself regarded perceptual judgments as "extreme cases" of abduction (CP 5.181). Other examples of abduction, also mentioned by Peirce (CP 2.714), include retroductive historical inferences. Peirce further pointed out that in science the abductive step is followed by severe observational and empirical tests of the deductive or probable consequences the hypothesis (CP 2.634; EP 2:114). The examples of abduction thus range from compelling everyday observations to the tentative adoption of theoretical hypotheses in science by virtue of their explanatory and predictive power. In these cases, it appears that abductive arguments can sometimes serve in providing a fallible *justification* of a hypothesis (see Niiniluoto, 1999b).

Peirce's own account of the *truth-frequency* of inference was later followed by many frequentist theories of probability and statistics in the 20th century. By this standard, the general reliability of abductive inference may be relatively high in some kinds of circumstances.

The *Bayesian* theory of inference uses epistemic probabilities: $P(H/E)$ is the rational degree of belief in the truth of hypothesis H given evidence E. The notion of confirmation is defined by the *Positive Relevance* criterion: E confirms H if and only if $P(H/E) > P(H)$. According to Bayes's Theorem, $P(H/E) = P(H)P(E/H)/P(E)$. If H logically entails E, we have $P(E/H) = 1$. Hence,

(2) If H logically entails E, and if $P(H) > 0$ and $P(E) < 1$, then $P(H/E) > P(H)$.

This result gives immediately a Bayesian justification for the Principle of Converse Entailment

(CE) If hypothesis H logically entails evidence E, then E confirms H

which has been taken to be characteristic to "abductive inference" (see Smokler, 1968; Niiniluoto & Tuomela, 1973). More generally, as positive relevance is a symmetric relation, it is sufficient for the confirmation of H by E that H is positively relevant to E. If inductive explanation is defined by the positive relevance condition, i.e., by requiring that $P(E/H) > P(E)$ (see Niiniluoto & Tuomela, 1973; Festa, 1999), then we have the general result:

(3) If H deductively or inductively explains E, then E confirms H.

The same principle holds for empirical predictions as well. Hence, by (2) and (3), *empirical success confirms the truth of a hypothesis*.

The notion of abductive confirmation is weak in the sense that the same evidence may confirm many alternative rival hypotheses. A confirmed hypothesis need not be rationally and tentatively acceptable on evidence. A stronger notion of inference is obtained if one of the rival hypotheses is the *best* explanation of the facts. The strongest justification is obtained if the hypothesis is the *only* available explanation of the known facts. In such cases, abduction might be formulated as a rule of *acceptance*. In 1965 Gilbert Harman formulated *inference to the best explanation* by the following rule:

(IBE) A hypothesis H may be inferred from evidence E when H is a better explanation of E than any other rival hypothesis.

Comparison with Peirce's schema (1) suggests the following version of IBE:

(IBE′) If hypothesis H is the best explanation of evidence E, then conclude for the time being that H is true.

Already in his early account of hypothetic or abductive inference, Peirce discussed *deductive* and *probabilistic* explanations. But one should also allow *approximate* explanations: H approximately explains E when it is possible to derive from hypothesis H a statement E′ which is close to E. Approximate explanation includes the problem of curve-fitting where the original observational data E is incompatible with the considered hypotheses H, so that $P(E/H) = 0$. For this case, the probabilistic link $P(E/H)$ between the explanans H and the explanadum E has to be replaced by a measure of *similarity* or *fit* between E and H (see Niiniluoto, 1999b). However, here the evidence may still indicate that the best hypothesis is *truthlike*. This principle might be called *inference* to the best approximate explanation:

(IBAE) If the best available explanation H of evidence E is approximate, conclude for the time being that H is truthlike.

If degrees of truthlikeness are introduced (see Niiniluoto, 1987), then there is a natural addition to IBAE: the greater the fit between H and E, the larger the degree of truthlikeness of H in the conclusion. A variant of IBAE could replace truthlikeness by the weaker notion of approximate truth:

(IBAE′) If the best available explanation H of evidence E is approximate, conclude for the time being that H is approximately true.

More technically, approximate truth can be defined by the minimum distance of the possibilities allowed by H from the truth, while the notion of truthlikeness combines the ideas of closeness to the truth and information about the truth (see below).

By combining the ideas in IBE′ and IBAE, *inference to the best theory* can be defined by

(IBT) If a theory has so far proven to be the best one among the available theories, then conclude for the time being that it is truthlike.

(See also Kuipers, 1999; 2000.)

As we have already seen, many attempts to defend scientific realism by the "no miracle argument" appeal to forms of abduction which conclude that successful scientific theories are approximately true (e.g., Putnam, Psillos). In other words, they involve something like the principle IBAE′ (but without making the notion of approximate truth precise).

To save the no miracle argument against the charges of circularity (Fine, 1986) and incoherence (van Fraassen, 1989), one needs to defend abduction in the form of IBAE, IBAE′, or IBT (cf. Niiniluoto, 2007a).

Let us first note that the Bayesian approach immediately shows that $P(H/E)$ may be close to 1 and $P(-H/E)$ close to 0, when H is the only explanation of E. Bas van Fraassen's (1989) objection that it is incoherent to give an extra bonus to a theory for its explanatory success fails, as such a bonus is not needed to prove (2) and (3). Van Fraassen's "bad lot argument" is not convincing as such, either (see Niiniluoto, 2004, p. 74), but it points out that sometimes the best available explanation is not yet "good enough" to be acceptable (cf. Lipton, 1991). Instead of accepting a weak theory, it may be rational to suspend judgment and continue searching for a better one. On the other hand, the same idea has been included in IBT in the phrase "for the time being": for a fallibilist realist, the attempt to improve our so far best theory is always a viable option which can be expected to be successful at least in the long run.

Van Fraassen's point that the Bayesian justification (2) does not hold for hypotheses H with zero probabilities is relevant – even though the decision to assign $P(H) = 0$ for *all* genuine theories postulating unobservable entities is a questionable form of scepticism. There are, indeed, interesting cases where $P(H) = 0$ seems reasonable: H is a sharp point hypothesis with a zero measure or H is known to be a false idealization. Hence, for such cases, new tools from the theory of truthlikeness are needed.

Laudan (1984) in his well-known "confutation of scientific realism" demanded the realists to show that there is an "upward path" from the empirical success of science to the approximate truth of theories – and then a "downward path" from approximate truth to empirical success. It is evident that any "upward" link has to be fallible and corrigible, given the correct core of Laudan's remark that there are non-referring and false but yet to some extent empirically successful theories in the history of science (see also Stanford, 2006). But this "meta-induction" need not lead to the pessimistic conclusion that all present and future theories are far from the truth, if we can argue that later more successful theories are progressively closer to the truth than earlier less successful ones.

In my own work, I have tried to reply to Laudan's challenge by using the concept of truthlikeness (see Niiniluoto, 1984, Ch. 7), i.e., by appealing to something like IBAE and by making it precise with my own account of truthlikeness and its estimation (see Niiniluoto, 1987). Kuipers (2000) also gives a reply to Laudan by his "downward" Success Theorem and "upward" Rule of Success.

The probabilistic account of IBE, given by the results (2) and (3), cannot be directly applied to our problem at hand. These results establish *a probabilistic link between explanatory power and truth*: posterior probability $P(H/E)$ is the rational degree of belief in the truth of H on the basis of E, and thereby confirmation, i.e., increase of probability by new evidence, means that we rationally become more certain of the truth of H than before. But a rule of the form IBAE needs a link between approximate explanation and truthlikeness. The notion of probability (at least alone) does not help us, since the approximate explanation of E by H allows that H is inconsistent with E, so that $P(E/H)$ and $P(H/E)$ are zero. In other words, while high posterior probability is an epistemic indicator of truth, we need corresponding indicators of truthlikeness.

One important approach to IBAE′ is to define the notion of *probable approximate truth* PA (see Niiniluoto, 1987, p. 280). Then probabilistic links between explanation and truth, like (2), induce probabilistic links between explanation and approximate truth as well. It is also possible that $PA(H) > 0$ even though $P(H) = 0$. This helps us to give a reply to van Fraassen's point about hypotheses with zero probability (see Niiniluoto, 1999a; cf. Festa, 1999). But this kind of result does not yet justify IBAE′, since here H is compatible with E.

Another challenge concerns the justification of IBAE and IBT. My own favorite method of connecting objective degrees of truthlikeness and epistemic matters is based on the idea of estimating verisimilitude by the expected degree of truthlikeness ver(H/E) (see Niiniluoto, 1987, p. 269). Let Tr(H, C^*) be the degree of truthlikeness of H relative to target C^*, where C^* the complete truth expressible in a given framework. Hypothesis H is itself a disjunction of complete theories C_i, $i \in I_H$. According to the min-sum measure, Tr(H, C^*) is a weighted average of the minimum distance of the disjuncts of H from C^* and the (normalized) sum of all distances of the disjuncts of H from C^*. The minimum distance alone defines the notion of *approximate truth*. When the target C^* is unknown, the *expected degree of verisimilitude* ver(H/E) of H given evidence E is obtained by going through all potential candidates C_i, $i \in I$, for C^* and by balancing the

likeness $\mathrm{Tr}(H, C_i)$ of H to C_i by the inductive probability $P(C_i/E)$ of C_i on E:

(4) $\mathrm{ver}(H/E) = \sum_{i \in I} P(C_i/E)\mathrm{Tr}(H, C_i)$.

The measure (4) of expected truthlikeness gives us an epistemic indicator of truthlikeness. Its definition includes epistemic probabilities but it is not identical with posterior probability. Indeed, $\mathrm{ver}(H/E)$ may be non-zero, and even high, when $P(H/E) = 0$. If $P(C'/E)$ approaches 1, when evidence E increases, then $\mathrm{ver}(H/E)$ approaches $\mathrm{Tr}(H, C')$.

In order to reply to Laudan's "upward" challenge, we can show that at least in some interesting situations a better approximate explanation of E has more expected verisimilitude given E. Further, the expected verisimilitude of H given E can be high, when H approximately explains E (see Niiniluoto, 2005). Thus, explanatory success gives us a rational warrant for making claims about truthlikeness. We can also study under what ideal conditions the estimated degree $\mathrm{ver}(H/E)$ equals the objective degree $\mathrm{Tr}(H, C^*)$ (see Niiniluoto, 2007b). Thereby the notion of expected truthlikeness, explicated by the function ver, provides *a fallible link from the approximate explanatory success of a theory to its truthlikeness*.

4. Explaining the success of science

To conclude, let us still consider the "downward" explanation of the empirical success of science by the truthlikeness of theories (cf. Niiniluoto, 1999a). For a scientific realist, the truth of a theory means that it gives a correct description of non-observable reality. This explains the success of the theory in describing observable phenomena and guiding our practical action: if theory H is true, then all empirical deductive consequences of H (if any) are also true. The same principle holds for approximate truth as well. For the min-sum measure of truthlikeness, the situation is more complicated, but in any case high truthlikeness of H gives a constraint to the approximate truth of deductive predictions from H. Further, high truthlikeness guarantees high "overall" empirical success.

According to Peirce's principle of pragmatism, the "rational purport" of a theory lies in its "conceivable bearing upon the conduct of life" (CP 5.413). That our best scientific theories should be used in human action and be pragmatically successful is thus an important Peircean idea. But if truth is *defined* by success, as in the so called pragmatist notion of

truth, then this idea cannot be turned into an explanation of the success of science – and thereby "the ultimate argument for scientific realism" cannot be formulated as an abductive inference to the best explanation. To see this, note that Arthur Fine (1986) has argued that in the explanatory schema

(5) Theory H is pragmatically successful, because H is true,

an instrumentalist or anti-realist can replace the realist notion of truth by the pragmatist notion of truth. But, as this pragmatist notion defines truth as pragmatic success, Fine's suggestion would turn schema (5) into a non-explanatory tautology. Similarly, if truth is replaced by van Fraassen's notion of empirical adequacy, schema (5) again fails to be explanatory, since then it would "explain" the empirical truth of the consequences of H by their empirical truth.

Gerald Doppelt (2005) claims that scientific realism must explain the "explanatory success" of science rather than its empirical adequacy. But clearly it is too much to demand with Doppelt that the truth of a theory alone would explain its "simplicity, consilience, intuitive plausibility, and unifying power". Such epistemic utilites may very well be additional desiderata that are independent of truth. For example, a tautology is certainly true, but it need not be simple, and it does not have any explanatory power. On the other hand, truthlikeness combines the ideas of truth and information, so that it helps to establish interesting links between the realist virtues of a theory and its explanatory and unifying power (cf. Niiniluoto, 2007).

Laudan and van Fraassen have also suggested that no explanation of the success of scientific theories is needed, since theories are selected for survival by their success (see van Fraassen, 1999). This evolutionary move is not convincing, either, since it fails to point out any characteristic permanent feature of our best theories (such as their truthlike correspondence to reality) which accounts for their ability to yield successful explanations and predictions. It is a different matter to describe the selection processes which give us empirically successful theories and to explain why such theories are (and continue to be) successful.

Non-scientific explanations of the success of science – e.g. appeal to miracles or God's will – are not acceptable. Therefore, we may conclude that scientific realism is the *only* explanation of the empirical success of science.

References

Doppelt, G. (2005). "Empirical Success or Explanatory Success: What Does Current Scientific Realism Need to Explain?" *Philosophy of Science*, 72, 1076–1087.

Festa, R. (1999). "Bayesian Confirmation." In M. C. Galavotti & A. Pagnini (Eds.), *Experience, Reality, and Scientific Explanation* (pp. 55–87). Dordrecht: Kluwer.

Fine, A. (1986). "Unnatural Attitudes: Realist and Instrumentalist Attachments to Science." *Mind*, 95, 149–179.

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.

Kuipers, T. (2000). *From Instrumentalism to Constructive Realism: On Some Relations between Confirmation, Empirical Progress, and Truth Approximation*. Dordrecht: Kluwer.

Laudan, L. (1984). *Science and Values*. Berkeley and Los Angeles: University of California Press.

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

Niiniluoto, I. (1984). *Is Science Progressive?* Dordrecht: D. Reidel.

Niiniluoto, I. (1987). *Truthlikeness*. Dordrecht: D. Reidel,

Niiniluoto, I. (1999a). *Critical Scientific Realism*. Oxford: Oxford University Press.

Niiniluoto, I. (1999b). "Defending Abduction." *Philosophy of Science* (Proceedings), 66, S436–S451.

Niiniluoto, I. (2004). "Truth-seeking by Abduction." In F. Stadler (Ed.), *Induction and Deduction in the Sciences* (pp. 57–82). Dordrecht: Kluwer.

Niiniluoto, I. (2005). "Abduction and Truthlikeness." In R. Festa, A. Aliseda, & J. Peijnenburg (Eds.), *Confirmation, Empirical Progress, and Truth Approximation* (pp. 255–275). Amsterdam: Rodopi.

Niiniluoto, I. (2007a). "Abduction and Scientic Realism." In F. Keskin (Ed.), *The Proceedings of the Twenty-First World Congress of Philosophy*, vol. 12 (pp. 137–142). Ankara: Philosophical Society of Turkey.

Niiniluoto, I. (2007b). "Evaluation of Theories." In T. Kuipers (Ed.), *Handbook of Philosophy of Science: General Philosophy of Science – Focal Issues* (pp. 175–217). Amsterdam: Elsevier.

Niiniluoto, I. & Tuomela, R. (1973). *Theoretical Concepts and Hypothetico-Inductive Inference.* Dordrecht: D. Reidel.

Pihlström, S. (1996). *Structuring the World: The Issue of Realism and the Nature of Ontological Problems in Classical and Contemporary Pragmatism.* Acta Philosophica Fennica 59. Helsinki: Societas Philosophica Fennica.

Pihlström, S. (2008). *"The Trail of the Human Serpent is over Everything": Jamesian Perspectives on Mind, World, and Religion*. Lanham: University Press of America.

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

Putnam, H. (1997). "James's Theory of Truth." In R. A. Putnam (Ed.), *The Cambridge Companion to William James* (pp. 166–185). Cambridge: Cambridge University Press.

Smokler, H. (1968). "Conflicting Conceptions of Confirmation." *The Journal of Philosophy*, 65, 300–312.

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

van Fraassen, B. (1989). *Laws and Symmetry*. Oxford: Oxford University Press.