

Defending Abduction

Ilkka Niiniluoto[†]

University of Helsinki

Charles S. Peirce argued that, besides deduction and induction, there is a third mode of inference which he called “hypothesis” or “abduction.” He characterized abduction as reasoning “from effect to cause,” and as “the operation of adopting an explanatory hypothesis.” Peirce’s ideas about abduction, which are related also to historically earlier accounts of heuristic reasoning (the method of analysis), have been seen as providing a logic of scientific discovery. Alternatively, abduction is interpreted as giving reasons for pursuing a hypothesis. Inference to the best explanation (IBE) has also been regarded as an important mode of justification, both in everyday life, detective stories, and science. In particular, scientific realism has been defended by an abductive no-miracle argument (Smart, Putnam, Boyd), while the critics of realism have attempted to show that this appeal to abduction is question-begging, circular, or incoherent (Fine, Laudan, van Fraassen). This paper approaches these issues by distinguishing weaker and stronger forms of abduction, and by showing how these types of inferences can be given Peircean and Bayesian probabilistic reconstructions.

1. Peirce on Abduction. Charles S. Peirce made a distinction between three modes of scientific reasoning already in his Harvard Lectures during the spring of 1865. Starting from Aristotle’s doctrine of induction as the inference of the major premise of a syllogism (of the first figure), Peirce observed that there is “a large class of reasonings” that are neither deductive nor inductive: reasoning *a posteriori* to a physical hypothesis, or inference of a cause from its effect (*W* 1: 180). This reasoning, which Peirce called Hypothesis, can be represented as the inference of the minor premise of a syllogism. Besides Aristotle, Peirce’s inspiration and the term ‘hypothesis’ came from the 1800 edition of Immanuel Kant’s *Logic* (*W* 1: 451). This classification of inferences was elaborated in Peirce’s Lowell Lectures in the fall of 1866, and published in the next year. In Peirce’s series for *Popular Scientific*

[†]Department of Philosophy, P.O. Box 24, 00014 University of Helsinki, Finland; e-mail: ilkka.niiniluoto@helsinki.fi.

Monthly ten years later, this distinction was presented in the article “Deduction, Induction, and Hypothesis” (1878).

Peirce illustrated the two ways of inverting deductive arguments by the following example with a Barbara syllogism (in the first figure). *Deduction* is an inference of a result from a rule and a case:

- (1) *Rule*. —All the beans from this bag are white.
- Case*. —These beans are from this bag.
- \therefore *Result*. —These beans are white.

Induction is the inference of the rule from the case and result:

- (2) These beans are from this bag.
- These beans are white.
- \therefore All the beans from this bag are white.

Hypothesis is the inference of the case from the rule and result:

- (3) All the beans from this bag are white.
- These beans are white.
- \therefore These beans are from this bag. (*W* 3: 325; *CP* 2.623)

A typical example of Hypothesis in the sense of (3) has thus the following logical form:

- (4) Given the law $\forall x(Fx \rightarrow Gx)$,
 from Ga infer Fa .

Already in the Harvard Lectures 1865, Peirce made it perfectly clear that Hypothesis is an *inference to an explanation*:

We find that light gives certain peculiar fringes. Required an explanation of the fact. We reflect that ether waves would give the same fringes. We have therefore only to suppose that light is ether waves and the marvel is explained. (*W* 1: 267)

In the Lowell Lectures 1866, Peirce says that Hypothesis is the inversion of the corresponding *explaining syllogism* which typically has the structure (1) (*W* 1: 428, 425, 440, 452). Here Peirce formulated the subsumption theory of explanation, earlier advocated by John Stuart Mill in the 1840s and later made precise by Carl G. Hempel in the 1940s (see Niiniluoto 1993).

Inspired by John Venn’s *The Logic of Chance* (1866), Peirce formulated induction, hypothesis, and analogy as probable arguments in 1867, where probability is interpreted as a *truth-frequency*, i.e., the proportion of cases in which an argument “carries truth with it.” In 1878, he formulated probabilistic versions of the Barbara syllogism (1) and its inversions by replacing the universal rule ‘All the beans from this

bag are white' with a statistical generalization 'Most of the beans in this bag are white' (CP 2.508–516, 2.627). In this situation, we might characterize Hypothesis as an inference from an effect to its probabilistic or statistical cause.

The article "A Theory of Probable Inference" (1883) gives several models of *probable* and *statistical deduction* from a statistical premise. Here 'deduction' means inference from a population to a sample, and is not necessarily truth-preserving. Peirce was again explicit in reasserting his earlier view that Inductions and Hypotheses, as the two forms of "ampliative reasoning," are "inferences from the conclusion and one premiss of a statistical syllogism to the other premiss," and in the case of Hypothesis, "this syllogism is called the *explanation*." Indeed, Peirce repeated, "we commonly say that the hypothesis is adopted *for the sake of the explanation*," and a statistical syllogism "may be conveniently termed the explanatory syllogism" (CP 2.716–717). Here Peirce anticipated Hempel's 1962 model of inductive-probabilistic explanation (see Niiniluoto 1993).

One of Peirce's formulations of Hypothesis in the 1883 article is given by:

- (5) a has the numerous marks G, G', G'', etc.
 b has the proportion r of the marks G, G', G'', etc.

Hence, probably and approximately, b has an r-likeness to a.

Here a has an *r-likeness* to b if the proportion of the shared properties of a and b is r; hence, $r = 1$ if a and b are completely similar. Thus, (5) is a probabilistic formulation of inference by *analogy*, based upon a kind of enumerative induction of characters or qualities instead of things (CP 2.706).¹

In his Cambridge Lectures of 1898, Peirce said that induction is "probable reasoning in the third figure"; this is what Aristotle called *epagoge* (Peirce 1992, 139).² Hypothesis, abduction, or retroduction, on the other hand, is "probable reasoning in the second figure." This is what Aristotle called *apagoge*—often understood as indirect proof. Its general form is

- (6) If μ were true, π , π' , π'' would follow as miscellaneous consequences. But π , π' , π'' are in fact true.
 \therefore Provisionally, we may suppose that μ is true.

1. This idea of analogy was anticipated by J.S. Mill; see Niiniluoto 1988. Levi (1995, 85) suggests that this inference is identical to what Peirce called "qualitative induction" in 1905.

2. In recent work in Artificial Intelligence, abduction has been generalized to pictorial representations, i.e., inferences to visual non-sentential hypotheses that are able to "produce" given data (see Thagard and Shelley 1997). For abduction and AI, see also Aliceda-Llera 1997.

(Peirce 1992, 140.) It is also called “adopting a hypothesis for the sake of its explanation of known facts.”

In his papers in 1901–03, Peirce defined *induction* in a new way as “the operation of testing a hypothesis by experiment” (CP 6.526). This is in harmony with the hypothetico-deductive (HD) model of scientific inference, but Peirce did not restrict his interest to cases where the test evidence is logically entailed by the hypothesis (cf. Niiniluoto and Tuomela 1973), and did not deny that there is a process of reasoning to a hypothesis. By *abduction* or *retroduction* he meant 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). Thus abduction consists in “studying facts and devising a theory to explain them” (CP 5.145). The general form of this “operation of adopting an explanatory hypothesis” is this:

- (7) 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). This schema shows how a hypothesis can be “abductively conjectured” if it accounts “for the facts or some of them”.

Schema (7) is more general than the inference (4), since it is not any more restricted to any simple syllogistic form.³ Further, here A might be a general theory (or, in the earlier terminology, it might include both a “Rule” and a “Case”). Moreover, the conclusion is not A itself, but the assertion that “there is reason to suspect that A is true.” Indeed, as Peirce wrote in 1910, from an actual fact retroduction “only infers a *may-be*” (CP 8.238).

2. Abduction and Discovery. Peirce himself stated in 1883 that the scientific task of “discovering laws” is accomplished by induction, and “discovering causes” by hypothetic inference (CP 2.713). The idea of an inference from effects to their causes is in fact a part of Aristotle’s doctrine in *Analytica Posteriora* (see I, 13, 34). Medieval and Renaissance followers of Aristotle called it *resolutio*, in contrast to *compositio* which proceeds from causes to effects. These Latin terms are translations of the Greek terms *analysis* and *synthesis*. Sometimes resolution was thought to proceed by syllogisms, but some commentators iden-

3. According to Whewell (1847), induction involves discovery: the given facts are first discovered to satisfy some new conception π , and only then this “superinduced” concept is used as a basis of inductive generalization (see Snyder 1997). Peirce, on the other hand, argues that in a valid induction the character π has to be “predesignated” rather than “suggested by the examination of the sample itself,” so that “induction can never make a first suggestion” (Peirce 1992, 139).

tified it with induction. Galileo and Newton compared the method of natural science with the method of analysis and synthesis of the Greek geometers, which is perhaps the most powerful idea of heuristics in the history of science (see Hintikka and Remes 1974).

From this perspective, Peirce's Hypothesis (4) or abduction (7) corresponds to what Hintikka calls the upward propositional interpretation of geometrical analysis (Hintikka and Remes 1974, 12). Peirce generalized this idea to probabilistic inference, and elsewhere his logical account of reasoning was also linked with the analysis-of-figures interpretation of analysis (cf. Hintikka 1997).

N. R. Hanson suggested in 1958 that Peirce's schema (7) could be understood as providing "a logic of discovery." The standard objection is that the hypothesis A occurs already in one of the premises, so that (7) cannot tell how A was discovered in the first place.⁴ A more interesting possibility is to modify (7) by replacing the particular hypothesis A with a class K of hypotheses (see Hanson 1961, Niiniluoto 1984, 35):

- (8) The surprising fact C is observed;
There is reason to suspect that some hypothesis of kind K explains C,
Hence, there is reason to suspect that some hypothesis of kind K is true.

This reasoning might be an important intermediate step in the discovery of a scientific theory A which is of kind K. Peirce's account of Kepler's investigation of the orbit of Mars, the "greatest piece of Retroductive reasoning ever performed," is an example of this (*CP* 1.71–74). Another example, which shows that the inference justifying the second premise of (8) may be even deductive, is Newton's famous derivation of the form of his inverse-square law of gravitation from "phenomena" (i.e., Kepler's laws).

Schema (8) is also clearly related to many later accounts of discovery: e.g., Herbert Simon's (1977) view of discovery as the heuristic search of the best solution from a selected domain and Thomas Nickles's (1981) constraint model.

A historically interesting application of abduction as a heuristic method can be found in classical detective stories, as shown by the semiotical and logical essays collected in Eco and Sebeok 1983.

4. Hintikka (1998, 519) proposes that abductive inferences could be characterized "as answers to inquirer's explicit or (usually) tacit question put to some definite source of answers (information)." This question-theoretic model of knowledge-seeking presupposes that the desired information already exists in some source (e.g., data base). Genuine discoveries, where such information is not yet available anywhere, might be perhaps characterized as answers to questions that the inquirer puts to himself or herself.

C. Auguste Dupin, the hero of Edgar Allan Poe's novels in the 1840s, employed a method of "ratiocination" or "analysis" which has the structure of retrodution. Similarly, the logic of the "deductions" of Sherlock Holmes is typically abductive.

3. Abduction and Pursuit. Hanson (1961) proposed that Peircean abduction gives reasons for *suggesting* a hypothesis, as distinguished from reasons for accepting a hypothesis. According to Isaac Levi (1979; 1995, 85), the task of abduction is to establish the "ultimate partition" of *potential answers* to a cognitive problem. A similar point has been made by saying that the conclusion of abduction states that there are reasons for *pursuing* a hypothesis (see the discussions in Nickles 1980) or finding a hypothesis *testworthy*. The ability of a hypothesis to give at least a potentially correct answer to a question, or to give a potential explanation of the facts known so far, is an important part of the *before-trial evaluation* of a hypothesis (see Niiniluoto 1984, 34–36).

According to this weak conception, abduction is at best a route for finding worthwhile hypotheses, but does not lend any support or justification to a hypothesis. To confirm a hypothesis or to make it acceptable, we have to put it into trials or tests by deriving observable predictions from it.

The distinction between before-trial and after-trial criteria in the evaluation of hypotheses is made in the classical formulations of the HD method. For example, William Whewell required that a hypothetical theory ought to "explain phenomena which we have observed," but also "foretel [sic] phenomena which have not yet been observed" and are "of a kind different from those which were contemplated in the formation of our hypothesis" (Whewell 1847, 62–65).

Peirce complained in his later work that before the twentieth century he "more or less mixed up Hypothesis and Induction" (CP 8.227; cf. Kapitan 1997, 493). This somewhat cryptic remark is explained in the same letter: Peirce tells that he had failed to sharply distinguish cases where (i) a hypothesis is tested by deriving from it miscellaneous observable consequences and then comparing these consequences to phenomena and (ii) the hypothesis (first adopted by retrodution) is further found to explain some already known facts (CP 8.231–232). The latter situation is today known as the "problem of old evidence." As Peirce's remarks in 1905 make clear, he became aware that some historically important theories have received their strongest support from the surprising finding that they are able to explain some already well-known phenomenon (CP 2.759).

These observations indicate that the distinction between the explanatory and predictive power of a theoretical hypothesis is not always

clear: Hempel proposed in the 1940s to combine these notions into one concept of the *systematic power* of a scientific theory (see Hempel 1965). It thus seems that abduction as a motive for pursuit cannot always be sharply distinguished from considerations of justification.

4. Inference to the Best Explanation. The Peircean schema (4) of abduction is not generally valid in the logical sense, since it is an instance of the well-known fallacy of affirming the consequent. However, Peirce was aware that in some cases inference from effect to cause may be irresistible or compelling: the abductive suggestion may come to us “like a flash.” According to Peirce, this happens in *perceptual judgments* which are “an extreme case of abductive inferences”—they are “absolutely beyond criticism” as they are not “fully conscious” (CP 5.181–185). Thus, unless we accept phenomenalism which denies the causal element in perception, we have to acknowledge that our everyday knowledge about the observational properties of ordinary physical objects (such as stones and tables) involves abduction.

Another important area of human knowledge which is based upon abduction is *history*. For example, the fact that Napoleon Bonaparte once lived is not any more “susceptible of direct observation,” but we believe it because “its effects (such as the histories, the monuments, etc.) are observed” (CP 2.714).

In these examples about perceptual and historical judgments, hypothetical inference for Peirce is not only a method of discovery but also a fallible way of *justifying* an explanation.⁵ Thus, in the strong interpretation, abduction is not only an inference to a potential explanation but to the *best explanation* (IBE, for short).⁶ According to Gilbert Harman 1965, IBE recommends the inference from evidence E to hypothesis H when H is a better explanation of E than any other rival hypothesis. For Harman, this IBE is at the same time the basic form of all inductive inference.

Hempel required in his early papers that an adequate explanation

5. In his later work, Peirce stated that there is “no probability” about abduction, as “no definite probability attaches even to the mode of inference” (Peirce 1992, 142), and “probability proper had nothing to do with the validity of Abduction, unless in a doubly indirect manner” (CP 2.103). This argument does not exclude the possibility that such probabilities may be determined, and even high, under some special conditions. It seems that Peirce is here thinking mainly about “theory” as a process where hypotheses are “tried” and “held provisionally as long as the facts will permit”, as distinguished from “practice” which involves belief and action (Peirce 1992, 142, 176–78).

6. One way of linking abduction to a weak notion of justification is proposed by Aliseda-Llera (1997), who formulates “abductive expansion” and “abductive revision” as models of belief revision.

should be *true* (see Hempel 1965, 246). For a scientific realist, this is a natural desideratum. For example, when we explain the death of a person, the best explanatory account should identify the *real cause* of the explanandum event. However, in science we usually cannot directly verify the truth of an explanatory hypothesis, but at best we can make indirect inferences and tests about its truth value. It would be ineffective to formulate IBE as the rule to infer the true explanation from the given data. Therefore, as Hempel recognized, we need also the concept of *potential* explanation, and then it is important to ask whether the ability of a hypothesis to serve as a potential explanation of known facts in any way indicates its truth.

The notion of “the best explanation” is of course ambiguous in several ways. One distinction, which is important but often ignored, is between *deductive*, *inductive-probabilistic*, and *approximate explanations*. More precisely, when a hypothesis *H* explains evidence *E* relative to background assumptions *B*, the following conditions should be distinguished:

- (a) *E* is deducible from *H* and *B*
- (b) *E* follows from *H* and *B* with the probability $P(E/H\&B) < 1$
- (c) *E'* which is close to *E* is deducible from *H* and *B*.

Case (a) is typical in the HD model of science: theory *H* is a universal generalization, and $P(E/H\&B) = 1$. Case (b) is typical in statistical inference. Case (c) includes the problem of curve-fitting where the original observational data *E* is incompatible with the considered hypotheses *H*, so that $P(E/H\&B) = 0$. The relativity of these conditions to the background knowledge *B* means that the converse abductive inferences from *E* to *H* are context-dependent (cf. Day and Kincaid 1994).

Another important distinction is between local and global theorizing (see Hintikka 1968). In *local* theorizing, we are interested in finding an explanation of a particular piece of data or evidence *E*, while in *global* theorizing evidence *E* is only a steppingstone to a more generally applicable theory.

Hintikka suggests that in the local case it is natural to look for a theory *H* that transmits a maximal amount of information about the given evidence *E*. If semantic information is defined by a logarithmic measure of “surprise value,” this demand leads to the principle of *Maximum Likelihood*:

- (9) Given evidence *E*, accept the theory *H* which maximizes the likelihood $P(E/H\&B)$.

This rule is a standard method in statistical estimation problems (case (b) above).

For the deductive case (a), the rule (9) does not help much, since it cannot distinguish between potential explanations that make the likelihood $P(E/H\&B)$ equal to one. Hence, the choice of the best explanation is usually taken to depend on other desiderata. But even here a difference in the rival hypotheses may be detected by enlargening the description of the evidence E . For example, suppose we are looking for an explanation of the death of a certain person, and there are available several possible causes H, H', \dots of her death. However, if the case is given a fuller description, including facts about her life and the symptoms in her body, it may be that many of the potential hypotheses fail to give a causal explanation any more. In the limit, it may happen that only *one* of the potential hypotheses is left, and this is then certainly the “best” explanation in this situation.

For the approximate case (c), the probability $P(E/H\&B)$ has to be replaced by a measure of similarity or fit between E and H . Here the evidence may still indicate that the best hypothesis is truthlike (see Niiniluoto 1994).

In global theorizing, we are trying to find a good hypothesis H which explains also other things than the particular data E . But, as noted by Popper, if the goodness of a hypothesis H is measured by its posterior probability $P(H/E)$ relative to the total evidence E , then logically weaker hypotheses are favored over stronger ones.

Here it may seem natural to take the amount of the substantial *information content* of H , measured by $\text{cont}(H) = 1 - P(H)$, as the epistemic utility to be maximized:

- (10) Given evidence E , accept the theory H which maximizes the information content $\text{cont}(H)$.

However, this rule has the undesirable feature that it favors a logically inconsistent theory H (with $P(H) = 0$ and $\text{cont}(H) = 1$). Moreover, it always prefers a logically stronger theory to a weaker one. It does not help at all to replace $\text{cont}(H)$ by the relative content of H given E (see Hintikka 1968) or by Hempel’s (1965) measure of the systematic power of H with respect to E , i.e., $\text{syst}(H,E) = P(\sim H/\sim E)$ (cf. Niiniluoto and Tuomela 1973, 66). Hence, the rule IBE leads to absurd consequences if the “best explanation” is simply characterized as the theory which has the maximal degree of information content or systematic power. It is an open question whether IBE can be saved from these troubles by joining it with simplicity considerations (cf. Whewell’s “consilience” in Thagard 1978, “explanatory unification” in Friedman 1974) or with additional penalties for “conceptual problems” (cf. “problem-solving ability” in Laudan 1977).

The situation changes decisively if information content and system-

atic power are treated as *truth-dependent* epistemic utilities, as shown by Isaac Levi, Jaakko Hintikka and Juhani Pietarinen (see Hintikka 1968, Niiniluoto 1990). For example, if the utility of accepting H is $\text{cont}(H)$ when H is true and $-\text{cont}(\sim H)$ when H is false, then—after the calculation of expected utilities—the rule (10) is replaced by

- (11) Given evidence E , accept the theory H which maximizes
 $P(H/E) - P(H)$.

The same result is obtained by choosing the two utilities as $\text{syst}(T, E)$ and $-\text{syst}(\sim H, E)$. According to this rule, the best hypothesis should have both a high posterior probability given evidence and a high information content. As posterior probability is equal to the expected truth-value of H given E , the result (11) is also obtained by choosing truth and information as the basic utilities. Thus, (11) is a way of balancing the demands of truth and information (systematic power) in scientific inquiry.

If hypothesis H is logically stronger than hypothesis H' , i.e. H entails H' , then we have $P(H) \leq P(H')$ and $\text{cont}(H) \geq \text{cont}(H')$. If, in addition, both H and H' entail a contingent evidence statement E , then we have $P(H/E) - P(H) \leq P(H'/E) - P(H')$. In this case (a), the rule (11) recommends that a logically weaker hypothesis is always better than a stronger one. However, in the case (b) where the hypotheses do not entail the evidence, this need not be the case.⁷ An example is sketched by Nelson (1996, 404). More generally, in Hintikka's logic of inductive generalization the difference in rule (11) is maximized by the logically strongest generalization compatible with a sufficiently large evidence.

5. Abduction, Confirmation, and Truth. Peter Lipton (1991) suggests that IBE should be understood as "Inference to the Loveliest Potential Explanation": a *lovely* explanation is one which offers a high degree of "potential understanding." Lipton points out that the loveliest explanation is not always the same as the "likeliest": a *likely* explanation is probably true relative to the total evidence. (Note that in this terminology the Maximum Likelihood principle (8) is a local loveliness condition.) But he argues that at least on some conditions the loveliest explanation is also the likeliest potential explanation. Lipton, who does not formalize his argument, is criticized by Eric Barnes (1995). A more technical argument to the effect that epistemically true theories and "best-explaining theories" coincide has been given by Raimo Tuomela (1985).

7. Peirce himself (1992, 193) says that the "commonest fallacy of retrodution" is "the idea that the [antecedently] most probable hypothesis is the best."

The question whether explanatory power is truth conducive resembles the debate whether simplicity is a sign of truth. Hans Reichenbach argued in the 1920s that “inductive simplicity” is an indicator of probable truth, but “descriptive simplicity” is only a matter of convenience (cf. Niiniluoto 1994). Bas van Fraassen (1989) makes a similar distinction between “confirmational” and “informational virtues” of theories, arguing that explanatory power belongs to the latter type.

I think the only way of resolving disputes of this kind is to try to give precise reconstructions of confirmation relations and inferences involving notions like simplicity or explanatory power. Three important approaches can be distinguished: (i) qualitative confirmation, (ii) frequentist probability, and (iii) Bayesian epistemic probability.

(i) Inspired by Hempel’s work on the *qualitative* concept of *confirmation*, Howard Smokler (1968) suggested that abductive inference, in contrast to enumerative induction, satisfies the following conditions:

- (CE) (Converse Entailment) If H entails E, then E confirms H
- (CC) (Converse Consequence) If E confirms H and K entails H, then E confirms K.

It has also been proposed that abduction is definable by conditions CE* and CC* which are obtained from CE and CC by replacing the notion of entailment by the stronger notion of deductive explanation (see Niiniluoto and Tuomela 1973, 226). For example, CE* can exclude cases where the entailment is too trivial or *ad hoc*. In his recent work, Laudan (1996) also argues that hypotheses are not confirmed by all of their observational consequences.

(ii) A natural strategy, followed already by Peirce, is to analyze abduction or IBE as a form of *probabilistic* reasoning. For Peirce, the probability of a mode of argument is its *truth-frequency*, i.e., its ability to yield true conclusions from true premises. He applied this successfully to many kinds of inductive inferences, but in a general abductive rule like (4) the relative frequency of attribute F in the class of Gs might be any number between zero and one. But, given some background assumptions, the “validity” (i.e., truth-frequency) of a hypothetic inference of form (4) might be high⁸; for example, perception in ordinary circumstances is reliable in this sense. The same is true of some of the more complex types of statistical and analogical abductive inferences that Peirce formulated. Peirce’s program of hypothetic inference is realized in statistics by the Neyman-Pearson tests of significance: the

8. For this reason, Fumerton (1980) argues that a typical abduction should be understood as an enthymeme where the suppressed premise would turn it into an inductive inference.

power of a test is defined as the probability of rejecting the null hypothesis, if it is true, so that one minus power is the Peircean truth-frequency associated with the test.

(iii) In the *Bayesian* school, Bayes's Theorem has traditionally been used for the calculation of "inverse probabilities", i.e., the probabilities of causes given effects. Bayesian confirmation theory (Stanley Jevons, J. M. Keynes) was conceived as an account of induction as converse deduction. Peirce was critical about such applications, but even a frequentist can use Bayes's Theorem if the prior probabilities of causes are given as relative frequencies in a population.

The definition of confirmation by the *Positive Relevance* criterion satisfies condition CE: assuming that $P(H) > 0$ and $P(E) < 1$, if H entails E , then $P(H/E) > P(H)$. Hence, CE^* is satisfied as well:

(CE^*) If H and E are contingent statements, and H deductively explains E , then E confirms H .

As positive relevance is a symmetric relation, even a weaker premise is sufficient in CE^* :

(CE^{**}) If H is a positively relevant inductive explanation of E , then E confirms H .

The Bayesian approach thus immediately justifies the idea that explanatory success is confirmatory. However, positive relevance does not satisfy the converse consequence condition CC.

Bas van Fraassen (1989) has presented technical objections to the Bayesian treatment of abduction. As IBE is always restricted to a set of historically given or formulated hypotheses, it may lead to "the best of a bad lot." How could we know that the true hypothesis is among the so far proposed? And as this set is any case very large, its "random member" must be quite improbable. It has been pointed out that this criticism, if valid, would create serious troubles to van Fraassen's own "constructive empiricism" (Psillos 1996). In the Bayesian framework these worries have a straightforward answer: in order to apply Bayes's Theorem, the cognitive problem has to be treated in terms of a set of mutually exclusive and jointly exhaustive hypotheses—usually one of them is the "catchall hypothesis." In the simplest case, this set contains only two elements, a hypothesis h and its negation $\sim h$. So trivially one and only one element of this set is true. Results of the form CE^* show that hypotheses may receive support from their empirical successes, so that their posterior probability may be high.

Results about positive *confirmation* do not yet guarantee that H would be *acceptable* on evidence E . It may happen that no one of the rival hypotheses is clearly superior to others, but then suspension of

judgment is the most rational conclusion. To improve the situation, new evidence is needed, or alternatively the problem set has to be expanded by the introduction of new concepts. The extended cognitive problem contains then new rival explanations which may be evaluated by IBE.

Van Fraassen (1989) argues further that the attempt to give an extra bonus to a hypothesis on the basis of its explanatory power would lead to an incoherent probability function (for criticism, see Kvanvig 1994). But this is not the correct Bayesian way of giving credit for explanatory success. For cases where the theory entails the evidence such “plausibility” considerations (simplicity, systematic power, coherence with accepted background theories) are usually built into the prior probability $P(H)$ of H (see Salmon 1990). But, if H does not entail E , it may also happen that the explanatory success of H relative to E manifests itself in the rise of the posterior probability $P(H/E)$. (For similar accounts of “inductive simplicity,” see Niiniluoto 1994.) We have also seen in Section 5 that IBE can be construed as a reasonable inductive rule of acceptance by choosing systematic power as a truth-dependent epistemic utility (without presupposing that truth can be known for certain).

6. Conclusion: The Necessity of Abduction. What is the best way of arguing that our abductive inferential practices, though fallible, are and have been to some extent truth conducive? Just to postulate an inborn human capacity to hit upon true hypotheses, like Galileo’s *il lumen naturale*, is hardly convincing. It is more promising to argue that abductive inferences (in particular, our perceptual judgments) are largely reliable within our everyday life, and that this fact about the human species can be given a naturalistic evolutionary explanation: “All human knowledge, up to the highest flights of science, is but the development of our inborn animal instincts,” as Peirce put it (*CP* 2.754).

Another important argument concerning science is based on abductive inference from the empirical success of scientific theories to the explanatory hypothesis that these theories are at least to some degree truthlike or approximately true descriptions of reality (see Smart 1989, Putnam 1978, Boyd 1984, Niiniluoto 1984, Leplin 1987, Lipton 1991). Here the abduction seems to be particularly strong, since all the proposed non-realist explanations of such empirical success seem to fail. In particular, the attempt to use the empirical adequacy of a theory T (i.e., the fact that all the empirical consequences of T are true) as the explanation of its empirical success (see Fine 1984) leads to a trivial non-explanatory tautology, if empirical success means the ability of T to yield true empirical consequences (see Leplin 1987, Niiniluoto 1990).

It is difficult to see how there could be any “nomological” link here, as André Kukla (1996) suggests. And if there were, would its existence be established by abduction?

Several critics have argued that the abductive defence of scientific realism is question-begging, since the whole issue about realism concerns the cogency of abduction (see Laudan 1984, 134; van Fraassen 1980, 1989; Fine 1984, 115). The realist may point out that abduction is already needed to establish the empirical adequacy of theories on the basis of their observed success so far (cf. Psillos 1996). As most of the critics of realism today are not advocating idealist or phenomenalist construals of human knowledge, it is important to add that already evidence statements about observations and measurements, as well as all statements about history, are based upon abduction. The standard statistical tests of significance also involve abduction. So clearly the critic of realism has the burden of proof in showing why abduction may be employed in some areas of science but not in others (see Boyd 1984, 67).

While these remarks concern the internal incoherence of non-realist attacks on abduction, a direct defence should follow Peirce's steps in the attempt to analyse the structure of abductive inferences in everyday life and science, and to find out the conditions of their probabilistic reliability.

REFERENCES

- Aliseda-Llera, Atocha (1997), *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. ILLC Dissertation Series, Amsterdam: University of Amsterdam.
- Barnes, Eric (1995), “Inference to the Loveliest Explanation”, *Synthese* 103: 251–278.
- Boyd, Richard (1984), “The Current Status of Scientific Realism”, in J. Leplin (ed.), *Scientific Realism*. Berkeley: University of California Press, 41–82.
- Day, Timothy and Harold Kincaid (1994), “Putting Inference to the Best Explanation in its Place”, *Synthese* 98: 271–295.
- Eco, Umberto and Thomas A. Sebeok (eds.) (1983), *The Sign of Three: Dupin, Holmes, Peirce*. Bloomington: Indiana University Press.
- Fine, Arthur (1986), *The Shaky Game: Einstein, Realism and the Quantum Theory*. Chicago: The University of Chicago Press.
- Friedman, Michael (1974), “Explanation and Scientific Understanding”, *The Journal of Philosophy* 71: 1–19.
- Fumerton, R. A. (1980), “Induction and Reasoning to the Best Explanation”, *Philosophy of Science* 47: 589–600.
- Hanson, Norwood Russell (1961), “Is There a Logic of Discovery?”, in H. Feigl and G. Maxwell (eds.), *Current Issues in the Philosophy of Science*. New York: Holt, Rinehart, and Winston, 20–35.
- Harman, Gilbert (1965), “Inference to the Best Explanation”, *The Philosophical Review* 74: 88–95.
- Hempel, Carl G. (1965), *Aspects of Scientific Explanation*. New York: The Free Press.
- Hintikka, Jaakko (1968), “The Varieties of Information and Scientific Explanation”, in B. van Rootselaar and J.F. Staal (eds.), *Logic, Methodology, and Philosophy of Science III*. Amsterdam: North-Holland, 151–171.

- . (1997), "The Place of C.S. Peirce in the History of Logical Theory", in *Lingua Universalis vs. Calculus Ratiocinator. Selected Papers 2*. Dordrecht: Kluwer, 140–161.
- . (1998), "What is Abduction? The Fundamental Problem of Contemporary Epistemology", *Transactions of the Charles S. Peirce Society* 34: 503–533.
- Hintikka, Jaakko and Unto Remes (1974), *The Method of Analysis: Its Geometrical Origin and Its General Significance*. Dordrecht: D. Reidel.
- Kapitan, Tomis (1997), "Peirce and the Structure of Abductive Inference", in N. Houser, D. D. Roberts, and J. Van Evra (eds.), *Studies in the Logic of Charles Sanders Peirce*. Bloomington: Indiana University Press, 477–496.
- Kvanvig, Jonathan L. (1994), "A Critique of van Fraassen's Voluntaristic Epistemology", *Synthese* 98: 325–348.
- Kukla, André (1996), "Antirealist Explanations of the Success of Science", *Philosophy of Science* 63 (Proceedings):S298–S305.
- Laudan, Larry (1977), *Progress and Its Problems*. Routledge and Kegan Paul: London.
- . (1984), *Science and Values: The Aims of Science and Their Role in Scientific Debate*. Berkeley: University of California Press.
- Leplin, Jarrett (1987), "Surrealism", *Mind* 96: 519–524.
- Levi, Isaac (1979), "Abduction and Demands of Information", in I. Niiniluoto and R. Tuomela (eds.), *The Logic and Epistemology of Scientific Change*. Acta Philosophica Fennica 30. Amsterdam: North-Holland, 405–429.
- . (1995), "Induction According to Peirce", in K. L. Ketner (ed.), *Peirce and Contemporary Thought*. New York: Fordham University Press, 59–93.
- Lipton, Peter (1991), *Inference to the Best Explanation*. London: Routledge.
- Nelson, David E. (1996), "Confirmation, Explanation, and Logical Strength", *The British Journal for the Philosophy of Science* 47: 399–413.
- Nickles, Thomas (ed.) (1980), *Scientific Discovery, Logic, and Rationality*. Dordrecht: D. Reidel.
- . (1981), "What is a Problem that We May Solve It?", *Synthese* 47: 7–32.
- Niiniluoto, Ilkka (1984), *Is Science Progressive?* Dordrecht: D. Reidel.
- . (1988), "Analogy and Similarity in Scientific Reasoning", in D. Helman (ed.), *Analogical Reasoning*. Dordrecht: Kluwer, 271–298.
- . (1990), "Measuring the Success of Science", in A. Fine, M. Forbes, and L. Wessels (eds.), *PSA 1990*, vol. 1. East Lansing, MI: Philosophy of Science Association, 435–445.
- . (1993), "Peirce's Theory of Statistical Explanation", in E. C. Moore (ed.), *Charles S. Peirce and the Philosophy of Science*. Tuscaloosa: The University of Alabama Press, 186–207.
- . (1994), "Descriptive and Inductive Simplicity", in W. Salmon and G. Wolters (eds.), *Logic, Language, and the Structure of Scientific Theories*. Pittsburgh: University of Pittsburgh Press, 147–170.
- Niiniluoto, Ilkka and Raimo Tuomela (1973), *Theoretical Concepts and Hypothetico-Inductive Inference*. Dordrecht: D. Reidel.
- Peirce, Charles S. (1931–1935, 1958), *Collected Papers*, (volumes 1–6, ed. by C. Hartshorne and P. Weiss; 7–8, ed. by A. Burks.) Cambridge, MA: Harvard University Press.
- . (1982–), *Writings of Charles S. Peirce: A Chronological Edition*, 1– (ed. by M. Fisch et al.). Bloomington: Indiana University Press.
- . (1992), *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898* (ed. by K. L. Ketner). Cambridge, MA: Harvard University Press.
- Psillos, Stathis (1996), "On van Fraassen's Critique of Abductive Reasoning", *The Philosophical Quarterly* 46: 31–47.
- Putnam, Hilary (1978), *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
- Salmon, Wesley (1990), "Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes", in C. W. Savage (ed.), *Scientific Theories*. Minneapolis: University of Minnesota Press, 175–204.
- Simon, Herbert A. (1977), *Models of Discovery*. Dordrecht: D. Reidel.
- Smart, J. J. C. (1989), *Our Place in the Universe*. Oxford: Blackwell.
- Smokler, Howard (1968), "Conflicting Conceptions of Confirmation", *The Journal of Philosophy* 65: 300–312.

- Snyder, Laura J. (1997), "Discoverers' Induction", *Philosophy of Science* 64: 580–604.
- Thagard, Paul (1978), "The Best Explanation: Criteria for Theory Choice", *The Journal of Philosophy* 75: 76–92.
- Thagard, Paul and C. Shelley (1997), "Abductive Reasoning: Logic, Visual Thinking, and Coherence", in M. L. Dalla Chiara et al. (eds.), *Logic and Scientific Methods*. Dordrecht: Kluwer, 413–427.
- Tuomela, Raimo (1985), "Truth and Best Explanation", *Erkenntnis* 22: 271–299.
- van Fraassen, Bas (1980), *The Scientific Image*. Oxford: Oxford University Press.
- . (1989), *Laws and Symmetry*. Oxford: Oxford University Press.
- Whewell, William (1847), *The Philosophy of Inductive Sciences, Founded upon Their History*, 2nd ed. London: John Parker and Sons.