Peirce on the justification of abduction

Peirce on the Justification of Abduction

Francesco Bellucci
Ahti-Veikko Pietarinen

Abstract. What sort of justification can be claimed for abduction? In this paper we reconstruct Peirce's answer to this question. We show that in his early works on the logic of science Peirce provided an abductive justification of abduction, and that in his mature writings the early solution is enriched by a reference to the place that abduction has in a typical scientific inquiry. Since abduction is the first stage of inquiry by which a hypothesis is suggested and which then has to be subjected to inductive testing, the fundamental abduction or *ur*-abduction that justifies abduction has also to be subjected to a verification by means of a fundamental induction or *ur*-induction, namely the abduction that abduction is valid is verified by an appeal to the history of science.

Keywords. Peirce; abduction; hypothetic-deductive method; induction; logic; classification of arguments; scientific inquiry.

1. Introduction

What sort of justification can be claimed for abduction? According to Peirce, the question of the justification of abduction is the "bottom question of logical Critic" (EP 2, p. 443). Logical critic is the central and most fundamental department of logic, the task of which is to discuss "the justification of each of the essentially different kinds of reasoning" (R 852, p. 1, 1911). The essentially different kinds of reasoning are for Peirce deduction, induction, and abduction, and logical critic has to provide an explanation of the justification of each.

In recent years, research on abduction has centred either on the problem of the logical form of abduction (e.g., Kapitan 1997) or on the attempt to provide a taxonomy of abductive inferences (e.g., Eco 1983, Thagard 1988, Magnani 2001, Hoffmann 2011, Park 2017). The problem of the logical form of abduction was central to Peirce, while he was much less interested in the classification of abductive forms. Peirce thought that deduction is of two types ("theorematic" and "corollarial"), and that induction is of three types ("crude," "qualitative," and "quantitative"), but he never says, as far as textual evidence can be trusted, that there are different types of abduction. Thus it is somewhat paradoxical that in the literature so many attempts have been made to find and define sub-types of abduction. Of course, many of those attempts have proven to be useful developments and ameliorations of Peirce's original notion. But this does not change the fact that the "discoverer" of this form of reasoning thought it to be of one, general type only.

On the other hand, it has often been argued that there is no "logic" of abduction, because abduction concerns discovery (heuristic), not justification (logic). But in fact, the distinction between the "contexts of discovery" and the "contexts of justification" was in some sense endorsed by Peirce himself. If the distinction is taken to demarcate the generation of a new idea or hypothesis (discovery) from the defense or verification of it (justification), then Peirce was the first or among the firsts to see that a complete scientific inquiry begins with the generation of hypotheses (by abduction) which then have to be verified (by deduction and induction, i.e. by inductively verifying its deductive consequences). But the distinction between "contexts of discovery" and the "contexts of justification," as such, is of no help to the solution of the problem of the justification of abduction: for the bottom question of logical critic is not how to justify (i.e. verify) the hypotheses arrived at by abductive

reasoning, but how to justify abductive reasoning itself. The "justification" sought in logical critic is a justification for the validity of a principle of reasoning, not the justification for a scientific proposition (hypothesis) arrived at by means of reasoning.

K. T. Fann's little monograph on *Peirce's Theory of Abduction* (1970) is still to this day the only book-length treatise on this fundamental aspect of Peirce's logic.[1] Fann devotes an entire section of the monograph to the description of Peirce's several attempts to justify abduction, and concludes that those attempts remain unsatisfactory. In particular, having detected the necessary justificatory elements of abduction in Peirce's claim that the human mind bears an affinity to nature, Fann declares this justification itself an instance of an appeal to abduction, rendering it viciously circular and therefore unacceptable: "The affinity of mind with nature is an hypothesis which can only be arrived at by abduction and thus must not be used to support the validity of abduction" (1970, p. 54).

In this article we show that Fann's rejection of Peirce's affinity-thesis was too hasty and in fact based on a mistaken reconstruction of Peirce's own argument. We show that Peirce's mature idea of a justification of abduction is indeed circular, but it is not *viciously* circular. In particular, with the support of Peirce's perennial arguments from affinity that emerge from the late and unpublished manuscripts, the justification of abductive reasoning lies in the fundamental hypothesis, which we call the *ur*-abduction, that we have a power of truly knowing things by means of the resources that abductive reasoning provides. As with every abduction, such a *ur*-abduction must be verified, and since the only way of verifying an abduction is, on Peircean principles, by means of an appeal to inductive reasoning, the *ur*-abduction that we have a power of abduction has to be verified by an *ur*-induction. This explains, in addition to what the justificatory requirements for abduction are, Peirce's frequent recourse to arguments from the history of science to evidence what abduction is able accomplish. This is yet another aspect that the recent literature has tended to neglect, although arguments from the history of science are, we submit, a crucial element in Peirce's overall argument for the justification of abduction. Thus, if our reconstruction is correct, abduction for Peirce is directly justified through an *ur*-abduction, which in its turn is verified through an *ur*-induction. The justification of abduction is *directly abductive* and *indirectly inductive*.

The question of the justification of abduction is related to, but independent, of the question of scientific realism. The so-called "ultimate argument" or "no miracle argument" (Niiniluoto 2018, pp. 156–163), according to which the ability of scientific theories to explain facts and to give correct predictions would be a miracle unless those theories and predictions refer to real things and real laws, was first advocated by Peirce himself in "The Fixation of Belief" of 1877. In that seminal paper he argued that the scientific method of fixing beliefs is the only one that assumes a distinction between truth and falsity, and therefore the only one that assumes that there is a reality to which our belief correspond. In his words:

> Its fundamental hypothesis [of the scientific method of fixing beliefs], restated in more familiar language, is this: There are real things, whose characters are entirely independent of our opinions about them; those realities affect our senses according to regular laws, and, though our sensations are as different as our relations to the objects, yet, by taking advantage of the laws of perception, we can ascertain by reasoning how things really are, and any man, if he have sufficient experience and reason enough about it, will be led to the one true conclusion. The new conception here involved is that of reality. (W3: 254)

[1] Other works touching on the topic include Atkins (2016, p. 215).

If there were no reality, it would be a miracle that different men, starting from different empirical observations and applying different methodologies of investigation, would be led to the same representation of an object. Such "convergence" in the representation of the object is only explainable by the hypothesis that that object is real and that it really determines those different observations. The hypothesis of reality, thus, is the fundamental hypothesis of the scientific method.

This can be considered as a first formulation of the "no miracle argument", which as Niiniluoto argues, is itself abductive (2018, p. 157). However, the Peircean problem of the justification of abduction cannot be straightforwardly reduced to the problem of scientific realism, for the following very simple reason. The hypothesis of reality, itself the product of abduction, is also fundamental in *inductive* reasoning: the fact that successive samplings of a whole will yield increasingly correct representations of the whole is also based upon the assumption that the characters of the whole are *independent* of the characters of its samples, so that while each single sampling maybe unrepresentative, yet the very possibility of obtaining further samples guarantees that in the long run the whole will be adequately represented. Peirce was a realism (of the "scotist type," as he repeatedly says), but the question of the justification of abduction is not directly reducible to the question of reality. The hypothesis that there are real things is insufficient, according to Peirce, to guarantee the validity of abduction.

The paper is divided into three sections. Section 2 reconstructs Peirce's earliest statements of the argument for the justification of abduction as they appear in his 1865 Harvard Lectures on the logic of science. At that time, Peirce had considered deduction, induction and abduction (then termed "hypothesis") as distinct *forms* or *kinds* of reasoning. It is only much later that he came to see them also as three distinct logical *stages* of a typical process of scientific inquiry. We explain this in Section 3. Then, Section 4 illustrates Peirce's mature argument for the justification of abduction as it emerges from his published and unpublished writings from c.1907.

2. Peirce's early theory of abduction

Peirce's early theory of abduction has a decided semiotic flavor. The texts that provide his fullest exposition of that theory are what survives of the manuscripts of the Harvard Lectures of 1865 and of the Lowell Lectures of 1866. These have been printed in the first volume of Peirce's *Writings* (W1). After having introduced the two traditional logical quantities of "connotation" and "denotation" (also termed "comprehension" and "extension"), Peirce divides signs or representations into "copies" (later, "likenesses" and "icons"), which connote without denoting, "signs" in the strict sense (later, "indices"), which denote without connoting, and "symbols," which both connote and denote, and denote in consequence of the connotation (W1, p. 272). Usually, a combination of symbols is a symbol, but sometimes symbols combine in composite symbols that lose either their capacity of denoting or that of connoting. A composite symbol that has denotation without adequate connotation is an "enumerative term," and a composite symbol that has connotation but no denotation adequate to it is a "conjunctive term" (W1, pp. 278–279). A deduction is an inference through a symbol, induction an inference through an enumerative term, and abduction (at that time called "hypothesis") an inference through a conjunctive term (W2, pp. 446–447).

Here is an example:

MINOR PREMISE*: This is a horned, bearded, cloven-hoofed, ruminant animal.*
MAJOR PREMISE*: Goats are horned, bearded, cloven-hoofed, ruminant animals.*
CONCLUSION*: This is a goat.*

The predicate of both premises ("horned, bearded, cloven-hoofed, ruminant animal") is a conjunctive term: it connotes the characters severally connoted by the symbols of which it is composed, but it denotes nothing in consequence of that connotation: a goat is a horned, bearded, cloven-hoofed, ruminant animal, but so is a Siberian ibex. The connotation of the conjunctive term is unable adequately to fix its denotation. Thus, when we observe that this animal is horned, bearded, cloven-hoofed and ruminant, we abductively infer that it is a goat, because goats have those characters. We do this even though goats are not the only animals to have them. In semiotic terms, we turn a conjunctive term ("horned, bearded, cloven-hoofed, ruminant animal") into a genuine symbol ("goat"), and we do this because that symbol is one of the subjects of that conjunctive term. The leading principle of abduction is, in semiotic terms, that the symbol ("goat") that embodies a conjunctive term ("horned, bearded, cloven-hoofed, ruminant animal") is predicable of the same objects (this animal) as the conjunctive term is (W1, p. 186; R 839, p. 17).

Using "form" for "conjunctive term," Peirce stated this principle as "all forms are symbolizable" (W1, p. 282). In other words all conjunctive terms are replaceable by the symbols of which they are predicated. The question of the justification of abduction is thus the question of the justification of its leading principle. This justification, Peirce argues, can only take itself the form of an abductive argument. The reason is the following:

> [T]here are three distinct kinds of inference; inconvertible and different in their conception. There must, therefore, be three different principles to serve for their grounds. These three principles must also be indemonstrable; that is to say, each of them so far as it can be proved must be proved by means of that kind of inference of which it is the ground. For if the principle of either kind of inference were proved by another kind of inference, the former kind of inference would be reduced to the latter; and since the different kinds of inference are in all respects different this cannot be. (W1, p. 280)

The three kinds of reasoning (deduction, induction and abduction) are essentially distinct and irreducible. This is what after Kapitan (1997) and Hintikka (1998) is called the Autonomy Thesis (AT). Furthermore, the division into these three kinds of reasoning is exhaustive, that is, there is no reasoning of a fourth kind. Let us call this the Exhaustivity Thesis (ET). In the above passage Peirce argues that AT and ET entail that the justification of each leading principle of inference has to employ the very kind of reasoning that has thereby to be justified. If, for example, the justification of abduction were wholly deductive or wholly inductive, we would not be entitled to speak of irreducible kinds of inference, because the former would be wholly reducible to either of the latter.

To see this, it is useful to consider Peirce's work on the syllogistic figures, which in his early period he deemed to be crucial for the understanding of AT. Against Kant, who in *Die falsche Spitzfindigkeit der vier syllogistischen Figuren* had maintained that all syllogistic figures other than the first can be reduced to the first by means of immediate inferences, Peirce shows that the immediate inferences by means of which second- and third-figure syllogisms are reduce to first-figure syllogisms are themselves, when expressed syllogistically, of the very syllogistic figure that has

thereby to be reduced (W1, p. 514). That is, a second-figure syllogism can be reduced to a first-figure syllogism only by means of an immediate inference that has the syllogistic form of the second figure, and mutatis mutandis this is also true of the third figure. Each figure, then, is in this sense "irreducible" to the others, that is, reducible to them in the sense that it can be put in another figure, but not *wholly* reducible to that figure because the reduction employs the very figure that is so to be reduced.

Something similar happens with the three forms of reasoning, since in Peirce's early theory each corresponds to one of the figures of the syllogism (deduction corresponding to the first figure, abduction to the second, and induction to the third). If the justification of abduction were wholly deductive or wholly inductive, the peculiar "abductive elements" of abduction would be lost. On the contrary, abduction can be considered as an inversion of a deduction, or as the limiting case of a deduction, but this does not mean that its justification is wholly deductive. In fact, AT says that each form of reasoning is not wholly reducible to either of the others, i.e. that the justification of each cannot wholly consist in an appeal to either of the other kinds of reasoning.

AT thus implies that the justification of abduction has to be abductive. This will remain a constant element of Peirce's thought throughout his life. Fann was worried about circularity in Peirce's justificatory account of abduction. But now we are beginning to realize that such circularity lies at the heart of Peirce's account all the same. What has to be shown is that such circularity is not vicious.

Peirce's abductive ("hypothetical") argument for the justification of abduction ("hypothesis") is included in the second Harvard Lecture of 1865:

> To prove that all forms are symbolizable. Since this proposition relates to pure form it is sufficient to show that its consequences are true. Now the consequence will be that if a symbol of any object be given, but if this symbol does not adequately represent any form then another symbol more formal may always be substituted for it, or in other words as soon as we know what form it ought to symbolize the symbol may be so changed as to symbolize that form. But this process is a description of inference *à posteriori*. Thus in the example relating to light; the symbol of "giving such and such phenomena" which is altogether inadequate to express a form is replaced by "ether-waves" which is much more formal. The consequence then of the universal symbolization of forms is the inference *à posteriori*, and there is no truth or falsehood in the principle except what appears in the consequence. Hence, the consequence being valid, the principle must be accepted. (W1, p. 185)

This argument is supposed to prove abductively that the principle of abduction ("all forms are symbolizable," or "any conjunctive term can be substituted by the symbol of which it is predicated") is true. Now, abduction is inverse reasoning: inference of the antecedent of a conditional from its consequent (Niiniluoto 2011). It is, as Peirce also calls it, inference *à posteriori*: if the consequent (the "consequence" in the previous passage) is true, then the antecedent is true. If therefore we could determine the truth of the consequent of the principle of abduction, then that principle itself would have been proved true abductively. This is indeed Peirce's strategy in the argument, as he states that "since this proposition relates to pure form it is sufficient to show that its consequences are true." In other words, since the principle to be proved is the principle of abduction, it is sufficient to prove it by abduction.

But what are the consequents of the principle of abduction? They are the single abductions that one may perform. In the previous passage Peirce refers to an example that he had just introduced in the same lecture (W1, p. 180):

MINOR PREMISE*: Light gives such and such phenomena.*
MAJOR PREMISE*: Ether waves give such and such phenomena.*

CONCLUSION*: Light is ether waves.*

The predicate of the minor premise ("giving such and such phenomena") is a conjunctive term. Peirce calls it a "symbol that does not adequately represent any form," that is, a composite symbol that connotes without thereby denoting. In the conclusion this predicate is substituted by a genuine symbol ("ether waves," or a symbol "more formal"). This way, a form ("conjunctive term") is symbolized. Peirce says that "as soon as we know what form it ought to symbolize the symbol may be so changed as to symbolize that form." This means that as soon as we know that that form may be predicated of the symbol, as in the major premise, the symbol may be substituted for that form, as in the conclusion. If therefore the principle of abduction (namely that "all forms are symbolizable") is true, such substitution of a form with a symbol that symbolizes it will always be permitted. But such substitution is permitted; therefore, the principle is true. In other words, the consequent of the principle is itself abduction, and since we *can* make abductions, the principle of abduction must be accepted.

Let us point out what this argument is not. In the first place, it is not a genuine "proof" of abduction, or an argument that independently demonstrates the truth of the leading principle of abduction. Peirce is clear that, given AT and ET, no independent proof is possible in so far as the leading principles of inference are concerned. *A fortiori*, it is not an *inductive* argument, either. Peirce is not claiming that since some of the abductions that scientists perform are correct, the principle of abduction must be for such reasons true. What the argument establishes is just that since we *may* reason abductively, that is, since we may substitute a form with a symbol that symbolizes it, then it may be that the principle of abduction ("all forms are symbolizable") is true. As Peirce states the issue, the argument that proves the truth of the principle of abduction "became certain only by speaking of that which has no sense except when this principle is true" (W1, p. 283). Since it makes sense to symbolize a form by a symbol, that is, since it makes sense to reason abductively, then the principle of the general symbolizability of forms is true. This is established abductively, by reasoning from the consequent (that it makes sense to reason abductively) to the antecedent (that the principle of abduction is true).

The argument is of course circular. However, given AT and ET, *any* argument that would prove the validity of a principle of inference would be circular in this way. For AT implies that *if* a proof of one of the three kinds of reasoning is possible, *then* it must use the very same reasoning to be proved by means of it. Thus any such proof – if possible at all – would be circular. In the third section we will see how in his mature works on abduction Peirce will enrich the structure of this fundamental argument in order to formulate his paramount justification of abduction.

3. Abduction as the first stage of inquiry

Peirce's mature articulation of scientific discovery rekindles the three *kinds* of arguments as distinct but connected *stages* of the typical process of scientific investigation:

> There are three kinds of reasoning based upon as many utterly distinct purposes and principles. They are severally used in the three logical stages of research. Not every inquiry goes through all the stages; for inquiry is not so utterly unlike other undertakings that a person may not enter upon it and go out of it at intermediate states of it. But the performance of a full and typical inquiry will take place in three distinct acts. (R 756, p. 1 c. 1906)

The first stage of inquiry is abduction, by which a hypothesis is suggested which explains certain observed, surprising facts. The explanation consists in this, that the hypothesis would lead to the observed facts as necessary consequences of it. Peirce presented what subsequently became the "classic" or "canonical" formulation of the logical form of abduction during the seventh and last Harvard Lecture of 1903:

> 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. (EP2: 231 = CP 5.189)

The conditional proposition that occurs as major premise ("If A were true, C would be a matter of course") expresses the relation between *explanans* (A) and *explanandum* (C). The conclusion does not assert that A is true. It states that there is reason to suspect as much. At other places Peirce tells that the conclusion of abduction is an *interrogation* (CP 6.528, 1901), expressible in a sentence in the interrogative mood (EP2: 287, 1903). In an unsent draft letter to Victoria Welby he writes:

> [The] "interrogative mood" does not mean the mere idle entertainment of an idea. It means that it will be wise to go to some expense, dependent upon the advantage that would accrue from knowing that Any/Some S is M, provided that expense would render it safe to act on that assumption supposing it to be true. This is the kind of reasoning called reasoning from consequent to antecedent. […] Instead of "interrogatory," the mood of the conclusion might more accurately be called "investigand," and be expressed as follows: It is to be inquired whether A is not true. The reasoning might be called "Reasoning from Surprise to Inquiry" (Peirce to Welby, July 16, 1905, RL 463).

The conclusion of abduction advances a hypothesis not as true, nor as a mere idea, but as an idea worth investigating in order to determine its truth as the inquiry proceeds. Its grammatical expression is that of a sentence in the "investigand" mood, meaning that the propositional content of the sentence (the hypothesis) is qualified by a peculiar illocutionary force, which Peirce expressed as "*It is to be inquired whether* H is true or not." This illocutionary force is not the interrogatory force as in standard questions of language, but the investigand force that scientific hypotheses have. It communicates a certain urge that one could observe certain hypotheses to possess, so much so that those hypotheses ought to be subjected to further investigation.

Further investigation means testing, and a hypothesis can be tested only with reference to its experimental consequences and observations we can make on how the outcomes of those tests agree with our experience. The second step of inquiry is the tracing of necessary, or deductive, consequences from the hypothesis. These consequences are precise experimental predictions from the hypothesis, and they ought to be selected independently of whether or not they are known to be true. The requirement that the hypothesis be experimental is satisfied by following the pragmatic maxim: only those hypotheses are admissible which can be put to a well-defined experimental test. It needs also to be noted that the surprising fact which prompted the formation of the hypothesis in the first place (the "C" in the Harvard example) is also a necessary consequence of the hypothesis: the explanation consists precisely in the fact that the hypothesis would lead to the observed phenomena as its necessary consequence. Naturally, aside from the surprising facts many other necessary consequences can be drawn from the hypothesis, which is necessary if we are to test the hypothesis at all.

The third stage of inquiry, then, consists in the testing of the hypothesis through a testing of those precise predictions. This is induction. It consists in considering the

predictions from the hypothesis, calculating what conditions should be satisfied in order for those predictions to be fulfilled, causing those conditions to be satisfied by the experiment, and observing the results of the experiment. If the predictions are fulfilled, the hypothesis is inductively conferred a certain confidence value. The inductive character of the procedure of hypothesis verification consists in this, that the exact predictions tested are a *fair sample* of all the predictions from the same hypothesis that could be tested, and all reasoning from sample to whole is inductive. The third step of inquiry is thus generalization: what is found true of some predictions would be found true for all of them.

4. The mature theory

In his mature works on the logic of science Peirce recurrently returned to the problem of the justification of abduction. Occasionally, he was inclined to affirm that the justification of abduction is, at bottom, inductive. In a 1901 piece on Hume he for instance writes:

> We say that a supposed state of things would *explain* a surprising phenomenon, if the latter would be a necessary consequence of the former. That the supposition, or *hypothesis*, as it is called would explain the phenomenon makes us think that it is perhaps a fact. It need not be a fact. Until it has been systematically tested, it remains a mere guess. Experience shows, however, that skillful guesses have a tolerably fair chance of being true (R 873, p. 2)

Skillful guesses have been shown by experience to be correct above chance level. Appeal to "experience" means appeal to induction, namely appeal to our experience of past abductions by which we infer by induction that abductions generally are correct above chance. In Peirce's 1901 paper "On the Logic of Drawing History from Ancient Documents" such references to the use of induction in the justification of abduction are explicit: "it is a primary hypothesis underlying all abduction that the human mind is akin to the truth in the sense that in a finite number of guesses it will light upon the correct hypothesis. Now inductive experience supports that hypothesis in a remarkable measure" (EP2, p. 108 = CP 7.220). A similarly explicit reference to induction as a justificatory basis for abduction is found in the draft of the 1906 "Prolegomena to an Apology for Pragmaticism": "Abduction is no more nor less than guessing, a faculty attributed to Yankees. Such validity as this has consists in the generalization that no new truth is ever otherwise reached while some new truths are thus reached. This is a result of Induction" (R 293, 1906). The validity of abduction is in Peirce's view based upon two general propositions: that neither deduction nor induction can "start a new idea" (CP 2.96), and that abduction does sometimes start a new idea that subsequently is found to be true. Both these generalizations are the results of induction. That some of the scientists' abductions have been proven true is, Peirce later writes in a long letter to Carus, "the most surprising of all the wonders of the universe":

> As for the validity of the hypothesis, the retroduction, there seems at first to be no room at all for the question of what supports it, since from actual fact it only infers what *may be*, – *may be* and *may be not*. But there is a decided leaning to the affirmative side and the frequency with which that turns out to be an actual fact is to me quite the most surprising of all the wonders of the universe (Peirce to Carus, August 1910, CP 8.238; Peirce 2014, 282).

This can be seen as a variant of the so-called "no miracle argument" (Niiniluoto 2018, pp. 156–163), which in brief states that since the theories arrived at by means of

abduction have been proven frequently true, or true more often than not, if abduction did not possess any kind of validity or justification, the discovery and existence of those theories would really be a miracle. Niiniluoto justly observes that the no-miracle argument is itself abductive, an observation that was made by Peirce much before the argument had neither acquired a name of its own nor observed to be such by others:

> This mysterious convincingness which the history of science seems to show is well borne out by the relatively small proportion of good retroductions that have turned out to be quite false. This statistical argument, which, it must be noted is itself retroductive (so that we must be on our guard against a begging of the question), is one of the supports, though by no means the principal support, of my doctrine that the human mind has a power of divination.[2] (R 652, pp. 23–24, 1910)

This passage is remarkable for two reasons. The first is that, unlike in his previous accounts, Peirce is now able to claim that the statistical argument from the history of science is in fact abductive ("is itself retroductive"), not inductive. We can formalize such an argument as follows:

1. *It is observed that scientific theories obtained by abduction are relatively often true.*
2. *If abduction were a valid mode of inference, then the relative frequency of the truth of those theories would be a matter of course.*
3. *Hence, there is reason to think that abduction is a valid mode of reasoning.*

In principle, nothing would prevent us from formalizing the argument as an induction. But in the passage quoted Peirce claims that the "frequency argument," i.e. the inductive argument, while it is one of the supports for the justification of abduction, yet it is "by no means the principal support." The question thus arises as to what Peirce thought the principal support for abduction to be.

The manuscripts R 652 soon trails off, with any remaining pages they may have been lost, but in an undated manuscript dating c.1906 Peirce had written on the same question:

> This kind of reasoning [abduction] is justified by two propositions taken together. One is that man's mind which is a natural product formed under the influences which have developed Nature (here understood as including all that is artificial), has a natural tendency to think as Nature tends to be. This must be so if man is ever to attain any truth not directly given in perception; and that he is to attain some such truth he cannot consistently, nor at all, deny. The other proposition is that no other process of deriving one judgment from another can ever give any substantial addition to his knowledge; so that, if he is to reason at all, we must assume that this kind of reasoning succeeds often enough to make it worth while; since it certainly is not worth while to leave off reasoning altogether. (R 876 CSP 3, c. 1906)

In this passage the justification of abduction is said to repose on two propositions taken together. The first is that the human mind has a natural tendency to reason correctly about natural phenomena because it is a product of the same influences, forces and drivers as natural evolution is. This naturalistic argument for the justification of abduction can also be found in the Harvard Lectures of 1903 (EP2, pp.

[2] Lest it be misunderstood, by "divination" Peirce meant a specific form of instinctive reasoning akin to other, familiar types of instinct, explaining it in the same manuscripts such that "the human mind possesses, in some degree, a power of *divining* the truth, which is no more, at its utmost, than to have some endowment of instinct such as many species of birds, insects, and other creatures possess" (R 652 CSP 14, 1910).

217–218), in the 1908 "Neglected Argument for the Reality of God" (EP2, pp. 443–445), and in a fragment dating probably 1907 contained in R 319(a).[3] The second proposition is that no other kind of inference can "start a new idea" (CP 2.96). Peirce thus concludes from this second proposition that either we must reason abductively or we do not reason at all. In an incomplete draft letter for Lady Welby of 1905, the two "propositions" that in R 876 are said to collectively justify abduction are distinguished as the "cause" and the "reason" of the validity of abduction, respectively:

> The *cause* (not the *reason*) of the validity of this mode of reasoning must be that man's mind having been formed under natural influence has an indefinite tendency toward believing the truth. But the *reason* for trusting to this method of reasoning is the reason of despair; i.e. that is, we have either to [*unknown next page*] (Peirce to Welby, July 1905, RL 463, p. 51)

Since the follow-up pages have been lost, we can only guess that what Peirce meant by the reason for trusting abduction was that "we have either to despair of knowing things or to reason abductively." This is the disjunctive consequence of the second proposition of R 876. The trustworthiness of abduction grows out of the irritation of living doubt when it would cause thought to act merely out of desperation. Such irritation has to be appeased by the belief that we can have substantial additions to our knowledge. To do that, it is better to rely on the truth of our many guesses.[4]

In sum, the *cause* of the validity of abduction is that the human mind has developed, as an effect of natural evolution, an instinct for guessingly correctly, while the *reason* of it is that no other kind of reasoning would allow us to know nature, and thus we either reason abductively or stop reasoning altogether. Choosing the latter would mean the progress of science to cease, just as a game of Whist would, following one of Peirce's examples, stop in certain situations if it were not commonly

---

[3] We quote only the latter: "What is that power of guessing right? It is the most important ingredient of good sense. No matter how far it a strictly congenital heritage [sic.] nor how far it is transmitted by tradition and teaching, it certainly corresponds to what we call Instinct in animals. Young wrens have to be taught to fly and young ducks to swim by their mothers; but they could not be taught to do either unless their minds had a natural genius for swimming or flying. They naturally know what sort of efforts to make, although they do not know it so well as to be able to do it without teaching. If a chemist, on pouring something into a vessel containing other things, meets with a great surprise in a resulting phenomenon contrary to his previous experience, he sets about looking for the cause of it, first in the vessel, and then, perhaps, elsewhere in his laboratory. His power of guessing saves him from wasting time in conjectures that it is because some old woman in another part of the town has been sticking pins into a waxen figure or because Mars and Jupiter are at that time distant from one another by an aliquot part of 360 degrees; and this, unquestionably, a native, inborn agreement of his mind with nature, due to his mind having been developed under the powers of nature and in the way in which nature itself has been developed" (R 319(a), pp. 1–3).

[4] Peirce's letter to William James (Dec. 25, 1909, EP2, pp. 501–502) also alludes to the argument from desperation: "The first kind of warrant consists in the reasoner's being *disposed to believe* in his proposition. This goes toward warranting the belief, since the very undertaking to find out a truth one does not directly perceive assumes that things conform in a measure to what our reason thinks they should. In other words our Reason is akin to the Reason that governs the Universe (we must assume that or despair of finding out anything. Now, despair is always illogical),—and we are warranted in thinking so, since otherwise all reasoning will be in vain. If it be so, a strong inward impulse to Believe a given proposition, tends to show that proposition to be true." Much earlier, in (CP 1.405 = W6, p. 206, 1887–1888), he takes despair to be "insanity. True, there may be facts that will never get explained; but that any given fact is of the number, is what experience can never give us reason to think; far less can it show that any fact is of its own nature unintelligible. We must therefore be guided by the rule of hope, and consequently we must reject every philosophy or general conception of the universe, which could ever lead to the conclusion that any given general fact is an ultimate one."

accepted among the players that there can be such situations that fully warrant "a player for acting on the hopeful hypothesis" (R 652, p. 14).

Another incomplete draft manuscript of 1910 offers the same disjunctive reason for the validity of abduction: the justification of abduction is that abduction "is the result of a method that must lead to the truth if there it is possible to attain the truth. Namely we must assume the human mind has a power of divining the truth, since if not it is hopeless even *[*to reason*]*" (R 276, p. 9, 1910). In R 652, p. 14, the reason is expressed as "to say we really believe in the truth of any proposition is no more than to say we have a controlling disposition to behave as if it were true," while this belief is, in turn, "sufficient reason for believing that we have such power" of guessing right.

Proceeding in this manner one can unearth plenty of evidence in Peirce's writings for an inductive justification of abduction, at times also considered abductive. One can also find enough support for the distinction he makes, sometimes explicitly and sometimes implicitly, between the *cause* of the validity of abduction (the development of instinct) and the *reason* for it (the avoidance of despair). Yet, we believe that Peirce's strongest case for the justification of abduction is contained in the following passage from R 637, one of the drafts of the "Preface" that he intended to add to the re-publication of his *Illustrations of the Logic of Science*.[5] In this passage we find both the inductive element that we have expounded above and the reference to the whole process of inquiry of which abduction is the first stage. Here is the key passage, in which "retroduction" just means "abduction":

> The logical justification of Retroduction […] is as follows. In the first place, we certainly do thoroughly believe and cannot help so believing, do what we may, that some reasonings are sound. For we can free ourselves of a belief only by reasoning ourselves out of it, and to do this is to believe that some reasonings are sound. Now although it is, of course, one thing to believe a proposition, no matter how thoroughly and firmly, and quite another for the proposition to be true, yet practically for the believer they are one and the same. For if his belief is perfect he thinks he is sure it is true and between that and his thinking it is true there is no practical difference. We must and do admit, therefore[,] that some reasonings are sound. But to say this is to say that some instinct or natural impulse to believe is in conformity with the real nature of things; and the only question is how far that conformity extends. This can only be ascertained by sampling; and the process of sampling will consist in taking Retroduction after Retroduction and testing the truth of each by as large a sample of its consequences as can conveniently be obtained. This justifies Retroduction, which simply puts that process of testing into practice for single Retroductions; and there is nothing in the justification that cannot be learned from indubitable external observation and equally indubitable reasoning (R 637, pp. 13–14, October 1909 = Peirce 2014, pp. 250–251).

Let us attempt to spell out the details of this argument. The argument is based upon the premise that some reasonings are sound. This premise is indubitable. We actually have direct experience of sound inferences, and thus doubting that some reasoning is sound (say, by attacking valid inference), would amount to doubting something of which we have direct experience. But such a doubt would only be a *feigned* doubt, and although it is unscientific and unphilosophical to suppose that any particular fact will never be doubted, yet on Peircean principles "we cannot go behind what we are unable to doubt" (W3, p. 14). In logic we must begin with some beliefs which are not *de facto* put to doubt, even though they may be doubted in the future. Thus, logic requires that we admit that some reasonings must be sound.

This admission is equivalent to the admission that, in some cases at least, we can truly know things as they really are. In order to explain this we make the hypothesis that we truly have some power of knowing things, that we have a power of

[5] See Peirce (2014).

abduction. Peirce put this in terms of the hypothesis that we have "some instinct or natural impulse to believe […] in conformity with the real nature of things." As another passage from the 1908 "Neglected Argument" recites, "[t]here is a reason, an interpretation, a logic, in the course of scientific advance; and this indisputably proves […] that man's mind must have been attuned to the truth of things in order to discover what he has discovered. It is the very bed-rock of logical truth" (EP2, p. 444). We have seen above Peirce's recurrent claims that the human mind has evolved an instinct to correctly represent reality. In the context of his argument in R 637, this hypothesis becomes the fundamental hypothesis of scientific inquiry, "the bed-rock of logical truth." It is the *ur*-abduction and the totality of justification of abduction rests upon it.

Now, the *ur*-abduction that we have a power of abduction must, like every abduction, be put to test to see whether and to what extent it stands up to experimental probing. Once the existence of sound reasoning, namely reasoning by which we can truly know reality, is admitted, the only question that remains to be settled is *how far* that knowledge extends, that is, what the limits of the hypothesis may be. "The only question," Peirce comments, "is how far that conformity extends". This is the real innovation over Peirce' early abductive justification of abduction (Section 2). In the early theory, abduction was simply justified by an abductive argument. In the mature theory, abduction is justified by an abductive argument (*ur*-abduction) *which has then to be verified*.

As we can gather from the previous discussion, a question concerning the truth of a hypothesis can only be answered by induction, i.e. hypotheses are verified inductively. Therefore, the extent to which we have a power of truly knowing real things by abductive reasoning is determined by checking what the reach of successful abductions is or what it has been in the past. This is where the history of science enters the game: since the history of science provides us with abundant examples of successful or partly successful abductive inferences, the fundamental abduction or *ur*-abduction that we have a power of abduction is confirmed, at least in some measure, by the results of the history of science. Peirce's account is as follows: "[S]ince all modern science depends ultimately on this method, its history furnishes such a sample of intelligent hypotheses, that a student of that history must be blind not to see that man's mind has a certain power of divining the truth" (R 638, pp. 14–15). The history of science provides as it were the material for making the *ur*-induction that verifies the *ur*-abduction that we have a power of abduction, just like single inductions verify single abductions. "This justifies Retroduction," Peirce concludes, as it "simply puts that process of testing into practice for single Retroductions."

If our reconstruction of Peirce's argument in R 637 is correct, then the justification of abductive reasoning lies in the fundamental hypothesis, or *ur*-abduction, that we have a power of truly knowing things by means of abductive reasoning. As with every abduction, this *ur*-abduction must be submitted to test, and it is tested through an argument from the history of science, which is our *ur*-induction. Abduction is therefore directly justified through an *ur*-abduction, which in its turn is checked inductively: the justification of abduction is for Peirce *directly abductive*, and *indirectly inductive*.

5. Conclusions

Why is this kind of justification circular but not viciously so? Peirce does not explicitly address this problem, but his answer comes clear from our exposition. On

the one hand, if one accepts AT and ET, the only justification that one can provide of abduction will itself be abductive. In this sense, the circularity of the justification of abduction is not vicious because it is legitimated by AT and ET: AT and ET allow any justification of abduction to be *legitimately* circular.

On the other hand, abduction is only the first stage of inquiry: it is nothing in itself and its entire value is assessed in relation to testing. In this sense, Peirce is not simply saying that abduction is valid because abduction is valid. He is saying that abduction is valid because the fundamental abduction that justifies it (the *ur*-abduction that we have a power of abduction) is itself verified by the history of science just as the first-level abductions are verified by experiment and induction in the course of typical scientific inquiries. Therefore, Peirce's justification of abduction is circular because it makes appeal to the very same kind reasoning that it justifies, but it is not *viciously* circular because the whole justificatory account goes through all the steps of a typical scientific inquiry. Thus the justification of abduction is not exhausted by that appeal, but in it recourse is made to the instruments of hypothesis testing and verification. The legitimate circularity of Peirce's justification of abduction derives from its being justified in two steps: directly by abduction, and indirectly by induction.

References

Atkins, R. K. (2016). *Peirce and the Conduct of Life*. Cambridge: Cambridge University Press.

Eco, U. (1983). Horns, Hooves, and Insteps. Some Hypotheses on Three Types of Abduction. In U. Eco & T. A. Sebeok (Eds.), *The Sign of Three*. Bloomington: Indiana University Press.

Fann, K. T. (1970). *Peirce's Theory of Abduction*. The Hague, Nijhoff.

Hoffmann, M. (2011). 'Theoric transformations' and a new classification of abductive inferences. *Transactions of the Charles S. Peirce Society* 46 (4), pp. 570–590.

Kapitan, T. (1997). Peirce and the structure of abductive inference. In N. Houser, D. D. Roberts, & J. van Evra (Eds.), *Studies in the Logic of Charles Peirce* (pp. 477–496). Bloomington: Indiana University Press.

Magnani, L. (2001). *Abduction, Reason, and Science: Processes of Discovery and Explanation*. New York: Kluwer.

Niiniluoto, I. (2001). Abduction, tomography, and other inverse problems. *Studies in History and Philosophy of Science Part A* 42 (1), 135–139.

Niiniluoto, I. (2018). *Truth-Seeking by Abduction*. Dordrecht: Springer.

Park, W. (2017). *Abduction in Context*. Dordrecht: Springer.

Peirce, Charles S. (1787–1951). A Harvard manuscripts (Charles S. Peirce Papers, MS Am 1632, Houghton Library, Harvard University) as listed in Richard Robin,

Annotated Catalogue of the Papers of Charles S. Peirce (Amherst, University of Massachusetts Press, 1967) (R); (RL) refers to letters that are listed in the correspondence section of Robin's catalogue.

Peirce, Charles S. (1932–1958). *The Collected Papers of Charles Sanders Peirce*. 8 vols, edited by C. Hartshorne, P. Weiss, and A. Burks, Cambridge, Harvard University Press (CP).

Peirce, Charles S. (1982–2009). *Writings of Charles S. Peirce: A Chronological Edition*, 7 vols, edited by. E. Moore, C. J. W. Kloesel et al., Bloomington, Indiana University Press (W).

Peirce, Charles S. (1998). *The Essential Peirce*. Vol. 2. Edited by the Peirce Edition Project. Bloomington: Indiana University Press (EP2).

Peirce, C.S. (2014). *Illustrations of the Logic of Science*. Edited by C. De Waal. Chicago: Open court.

Thagard, P. (1988). *Computational Philosophy of Science*. Cambridge, MA: MIT Press.