FOUR PROBLEMS OF ABDUCTION:
A BRIEF HISTORY

Anya Plutynski

Debates concerning the character, scope, and warrant of abductive inference have been active since Peirce first proposed that there was a third form of inference, distinct from induction and deduction. Abductive reasoning has been dubbed weak, incoherent, and even nonexistent. Part, at least, of the problem of articulating a clear sense of abductive inference is due to difficulty in interpreting Peirce. Part of the fault must lie with his critics, however. While this article will argue that Peirce indeed left a number of puzzles for interpreters, it will also contend that interpreters should be careful to distinguish discussion of the formal and strictly epistemic question of whether and how abduction is a sound form of inference from discussions of the practical goals of abduction, as Peirce understood them. This article will trace a history of critics and defenders of Peirce's notion of abduction and discuss how Peirce both fueled the confusion and in fact anticipated and responded to several recurring objections.

1. Introduction

Why should philosophers of science study their history? Popular but unsatis-fying reasons include that it permits us to learn from our mistakes or that we give credit where credit is due. Of course, we should credit the authors of important distinctions, arguments, or concepts, but we should also attempt to see through their eyes as a way of viewing our own problems and problem

Anya Plutynski is associate professor of philosophy in the Department of Philosophy, 215 S. Central Campus Dr., Rm. 402 CTIHB, Salt Lake City, UT 84112 (plutynski@philosophy.utah.edu).

space anew. An ongoing dialogue with its history should inform the philosophy of science today.

Peirce may seem an unlikely candidate for this ongoing conversation, for, by almost any measure, he was a bad conversationalist. He had a habit of making vague promises and leaving them unfulfilled; he was master of the tangential discussion; he changed the definitions of his key terms; he subscribed to genuinely bizarre metaphysics and epistemology. However, he also had keen insights, and reading his work is a helpful reminder that distinctions and norms that we take for granted as twenty-first-century philosophers of science are not givens.

Peirce was writing before the distinction between the context of discovery and the context of justification had been invented (or, perhaps, identified); he was writing before philosophers of science worried about whether their central task was descriptive or normative; he was writing before epistemology was "naturalized." These three differences between Peirce's context and that of philosophy of science today are instructive, particularly in the context of debates about abduction. Many, if not most, of the objections that have been raised against Peirce (and counterarguments in his defense) can be traced to competing views about the legitimacy of a demarcation between the context of discovery and that of justification, whether normative projects in philosophy of science can or should be addressed independently from descriptive accounts of scientific practice, and finally, what it could mean to be a "naturalist" about epistemology. Where you stand concerning these issues will, it turns out, have rather important implications for how you fall on the matter of abductive reasoning: its warrant, relevance, and import for science.

2. Origins: Abduction

Peirce first mentioned abduction, or what he called "reasoning by hypothesis," in 1866. He described this form of inference as from result and rule to case:

Light is polarizable. (Result; or, "Fact explained")
Ether waves are polarizable. (Rule)
Light is ether waves. (Case) (1982–2010, 1:427; hereafter WP followed by volume and page number)

Then, in 1878, he gives a similar illustration of inference to hypothesis:

Rule: All the beans in this bag are white.
Result: These beans are white.

228

Case: These beans are from this bag. (Peirce 1931–58, 2:623; hereafter
CP followed by volume and page number)

Kapitan (1997) paraphrases this:

All As which are B are C.
This A is C.
Therefore, the A is B. (480)

Abduction is contrasted with deduction and induction, which involve an
inference from either rule and case to result (All As which are B are C; This A is
B; Therefore, this A is C) or case and result to rule (This A is C; This A is B; All
As which are B are C), respectively. Peirce calls this third form of inference
"making an hypothesis" (1992–98, 1:188; hereafter EP followed by volume
and page number) and later "abduction" (WP 2:108), which he translates from
Aristotle's *apagoge* (*Prior Analytics* 2.25 69 20–36).[1]

Peirce insists, "All the ideas of science come to it by way of abduction." And
he roughly characterizes abduction as "studying facts and devising a theory
to explain them" (CP 5:145). Deduction is said to draw out consequences of
a rule and case; it is "necessary reasoning." Induction, in contrast, is "experi-
mental testing of a theory." And "hypothesis is where we find some very curious
circumstance, which would be explained by the supposition that it was the case
of a certain general rule, and thereupon adopt that supposition" (EP 1:189).

Peirce left a legacy of puzzles about abduction, yielding a contentious litera-
ture. Debates in the literature have often circled around one or more of four
central interpretive challenges, which I will call the formalization indeter-
minacy problem, the boundary problem, the justification problem, and the
descriptive problem. First, how (if at all) ought abductive inference to be for-
malized; second, what is the boundary between abduction and other forms of
inference; third, is abductive inference a sound form of inference, or how could
it be made so; and fourth, does abductive reasoning describe features of the
actual practice of science, and, if so, which features? Of course, these questions
are overlapping; how one formalizes abduction bears on whether it is a sound
form of inference, as well as whether one views it as distinct from other in-
ference forms. Moreover, all of these questions bear on how we characterize
abductive inference in scientific practice.

---

1. Peirce uses the term "retroduction" only after it is pointed out to him that his translation of
apagoge as abduction is problematic. For a detailed discussion of the differences and similarities
between Aristotle and Peirce's logic, see Kraus (2003).

A central problem with characterizing abduction, as Fann (1970) and Anderson (1986) have pointed out, is that Peirce subtly changes his characterization of abduction over the course of his career. Early in his career, Peirce places more emphasis on the formally distinct character and role of abductive inference. In 1866, it appears as the above simple-inference form. After the turn of the century, Peirce seems to place more emphasis on the role of abductive inference in scientific discovery, and he characterizes the logical form of the inference rather differently (see below). One of the major criticisms of abduction in the literature has been over reconciling these two characterizations. Some scholars have attempted to reconcile the nineteenth- and twentieth-century Peirce by arguing that he saw abduction as a two-step process: the generation and then the selection of hypotheses. Kapitan (1992) calls this "abductive discovery" and "abductive preference." This reconciliation helps make some sense of Peirce's ambiguity.

However, it also generates another set of problems. If abduction is a two-stage process, how are the two stages different? Do they use the same argument form or different ones? Are they guided by different logics? Is it even clear that one can demarcate such neat stages in the practice of science?

Another debate concerns whether Peirce thought of abduction as reason guided at all. In a 1903 paper, Peirce characterizes abduction as both clever guesswork—"the abductive suggestion comes to us like a flash. It is an act of insight" (CP 5:180–212)—and a logical form of reason-guided inference: "It must be remembered that abduction, although it is very little hampered by logical rules, nevertheless is a logical inference asserting its conclusion only problematically or conjecturally, it is true, but nevertheless having a perfectly definite logical form" (CP 5:188).

These apparently inconsistent claims have generated generations of debate, debate that has been closely tied to a shifting balance of views about the scope and character of philosophy of science, as a discipline. In 1958, Frankfurt asked the question, Is abduction an act of insight or an inference? Or, in his own words, Frankfurt states: "We are, then, faced with the seeming paradox that Peirce holds both that hypotheses are the products of a wonderful imaginative faculty in man and that they are products of a certain sort of logical inference" (1958a, 594). This makes the dichotomy seem rather starker than it need be since one might experience an inference as a "leap" of insight, without being aware of the logical steps of making an argument, but it is not surprising that a philosopher of science would ask the question just this way in the 1950s. Anderson, another interpreter of Peirce, argues that the two are not mutually exclusive: "Peirce does not hold insight and inference to be mutually

exclusive. … Abduction is both an insight and an inference. This is a fact to be explained, not to be explained away" (1986, 156).

Anderson's suggestion seems to be more than a mere claim about the nature of abduction: it is a call for historicizing and contextualizing philosophy of science. Anderson was writing in response to a long tradition that criticized Peirce's failure of precision, claiming that Peirce was interested not exclusively in the normative question of how discovery should go but also in the descriptive problem of how it does go. For Peirce, knowledge of the psychological character and the detailed context of scientific theory making and discovery were not irrelevant to a philosophical understanding of the growth of knowledge. Indeed, in some ways, Peirce here anticipates a contemporary turn in philosophy toward the social and the cognitive sciences. If, in fact, the process of abduction often occasions a feeling of "insight," perhaps looking to the sources and patterns of such feelings may tell us something interesting about the logic of science and not merely its phenomenology. Some psychologists have suggested that strong feelings often accompany the formation of convictions (Burton 2008). This does not, of course, guarantee the certainty of such convictions, but that there are such psychological features of discovery, and perhaps even identification of why and how these phenomenological experiences are instantiated or prompted, may lead us to learn something interesting about patterns of discovery, apart from their logic. Peirce was, arguably, ahead of his time in this regard. (For some contemporary examples of how philosophers of science have drawn on cognitive psychology, see, e.g., Gopnik and Schulz [2007].)

Next, I explicate four interpretive problems concerning Peirce's articulation of abduction. See table 1, for the four problems and for how different authors responded to these problems, discussed at greater length below.

2.1.   The Formalization Indeterminacy Problem

Peirce describes the inferential structure of abduction differently over the course of his career. For example, in the first paper where he mentions abduction, he describes the method of hypothesis as the third permutation of the syllogism mentioned above. In 1903, he describes the method of hypothesis as follows:

> The surprising fact, C, is observed.
> But, if H were true, C would be a matter of course, hence.
> There is reason to suspect that H is true. (EP 2:231)

Table 1. Competing Views on the Four Problems of Abduction

Author	Boundary Problem	Formalization Problem	Descriptive Problem	Justification Problem
Peirce (1982–2010), 1866, 1878, 1903, 1905	Acknowledges in 1878 but later insists that he confused "hypothesis" and "induction"	At least two formalizations: one is syllogistic, and one affirms consequent	Aims to be a description of how we arrive at novel hypotheses	Weakness of abduction in second formalization
Popper (1934/1968)	No distinct logic of discovery	Formalization impossible since no such form of inference	Discovery is a matter for psychology, not logic	Since no such inference, cannot be either weak or strong
Hanson (1958/1972)	Insists on firm boundary: abduction is a logic of discovery only	A, but A, would follow as a matter of course if B, therefore B	Description of discovery	Admits weakness
Frankfurt (1958a)	Abduction is vague/indistinct from induction	Formalization runs into difficulties; would involve Peirce in inconsistency	Given Peirce's formalization, abduction cannot be a method of discovery of novel hypotheses	Weak
Harman (1965)	No boundary: induction is a special case of abduction—both involve justification	Formalizes as relative likelihood	Aims to describe all of inference	Yes
Achinstein (1970, 1980, 1987)	There is a logic of discovery; abductive reasoning is not induction, but it is inference to the best explanation	Formalization impossible	Description of discovery not reducible to inferential form; depends on background knowledge, etc.	Weak, but influential and important in science
Thagard (1978)	There is a logic of discovery; it is inference to the best explanation	Formalization possible to some extent; however, context and historical features play a role	Possible to describe discovery using computer algorithms	Weak
Van Fraassen (1980)	Inference to the best explanation is incoherent	Formalization impossible; results in Dutch book	Descriptive problem is insoluble	Weak
Salmon (2001)	No boundary; inference to the best explanation is just Bayesian updating	Formalize abduction	Redescribe it as Bayesian	Reinterpreted as Bayesian updating, strong

This argument form has been challenged on a number of fronts. The hypothesis is included among the premises, so Frankfurt (1958a) argues that it cannot be that hypotheses are the result of abduction, as Peirce claims. Second, as a variety of critics have pointed out, the argument seems to permit too much. Indeed, Peirce himself anticipated this objection: "It may be said that the argument would prove too much; for if it were valid, it would follow that no hypothesis could be so fantastic as not to have presented itself" (CP 5:191). Any number of absurd hypotheses could potentially explain some fact; that a hypothesis could do so, however, is no reason to take it to be plausible. Various authors have attempted to add a third, "corrective" premise in light of this very objection.

However, Peirce anticipated this worry and elaborated his own set of criteria for assessing the relative merits of hypotheses. As early as 1878, Peirce suggested that there were "rules for the application of abduction, … in order that the process should lead to a probable result" (EP 1:193). He writes, "When we adopt a certain hypothesis, it is not alone because it will explain the observed facts, but also because the contrary hypothesis would probably lead to results contrary to those observed" (EP 1:191). In other words, Peirce anticipates an important insight about what happens after abduction: confirmation—the evaluation of hypotheses is always contrastive (see, e.g., Sober 1994). Of course, more needs to be said about what one decides to include in the contrast class and how to sort among the alternatives, as discussed below.

2.2. The Boundary Problem

The formalization indeterminacy problem is closely tied to the boundary problem: how to distinguish abduction from other forms of inference. It is unclear what falls under the umbrella of abduction. In particular, one puzzle that has particularly troubled interpreters is whether—and if so, how—abduction is distinct from induction. Peirce, in 1878, describes induction very narrowly as what some have called the "straight rule," or "induction by enumeration." He writes that induction merely "classifies," while abduction "explains" (EP 1:194), and he credits abduction with the arrival at general theories, as opposed to mere classification. Hypothesis, he explains, is when "we stretch induction quite beyond the limits of our observation" (EP 1:197). The method of hypothesis, he writes, allows us to arrive at knowledge of "causes and forces" and "enables us to see the why of things," whereas induction is merely "the process by which we find the general characters of classes and establish natural classifications" (WP 1:428). Peirce even categorizes the sciences in terms of which form of inference predominates. "Merely classificatory sciences" are

purely inductive—systematic botany, zoology, mineralogy and chemistry—whereas geology and biology (largely historical sciences) are "sciences of hypothesis" (EP 1:199). Despite his insistence that induction and abduction are distinct, early on, Peirce also says that the boundary between the two modes of inference can be difficult to discern: "The analogy of hypothesis with induction is so strong that some logicians confound them" (EP 1:192). To complicate matters further, later in his career, Peirce claims that he confused induction and abduction before 1900. One central puzzle in the literature on abduction, thus, has concerned the boundary between induction and abduction.

2.3. The Justification Problem

The formalization and boundary problem are closely tied to the justification problem: Is abduction a good form of inference? Some (again, Frankfurt 1958b) argue that the argument form affirms the consequent. That is, if the argument is "If H, then C is true. C is true. Therefore H," then the argument affirms the consequent. Peirce anticipates and replies to the objection as follows: "It is to be remarked that it is only in deduction that there is no difference between a valid argument and a strong one. … An argument is none the less logical for being weak, provided it does not pretend to strength that it does not possess. … An argument is fallacious only so far as it is mistakenly, though not illogically, inferred to have professed what it did not perform" (CP 5:192). Peirce here is suggesting that abduction does affirm the consequent, but abductive reasoning is not deductive reasoning and so is not subject to the same standards. Indeed, different argument types—inductive, deductive, and abductive—have different goals and thus should be judged by different standards of adequacy. Abduction is not to be judged on the same merits as deduction. (For a very similar argument in the contemporary psychological literature, see, e.g., Hahn and Oaksford [2007].)

The above response leaves unanswered the following question: On what merits is abduction to be judged? What then makes an abductive argument a strong one? How, for instance, are we to avoid the "anything goes" fallacy mentioned above? In response to this criticism, Peirce wrote: "As to the … objection that every hypothesis, however fantastic, must have presented itself … I have only to say that this could only arise in a mind entirely unpracticed in the logic of relations, and apparently quite oblivious of any other mode of inference than abduction" (CP 5:193). (Alas, after a careful study of the logic of relations, this particular claim is not yet altogether clear.) However, the general intuition seems to be that to understand how to differentiate better from worse abductive inferences, one must first be practiced in other forms of inference and,

moreover, have deployed them properly in generating the appropriate contrasting hypotheses.

Peirce explains that a variety of considerations enter into choice of hypothesis: it must be capable of experimental test, it must explain the facts in question, and it must be "economical."

> Now economy … depends upon three kinds of factors, cost; the value of the thing proposed, in itself; and its effect upon other projects. … Under the head of value, we must place those considerations which tend toward an expectation that a given hypotheses may be true. These are of two kinds, the purely instinctive and the reasoned. … Hypotheses … [that] naturally recommend themselves to the mind … make up on us the impression of simplicity—which here means facility of comprehension by the human mind, —of aptness, of reasonableness, of good sense. From the instinctive we pass to the reasoned marks of truth in the hypotheses. Of course, if we know any positive facts which render a given hypothesis objectively probable, they recommend it for inductive testing. When this is not the case, when the hypothesis seems to us likely, or unlikely, this likelihood is an indication that the hypothesis accords or discords with our preconceived ideas; and since those ideas are presumably based upon some experience, it follows that, other things being equal, there will be, in the long run, some economy in giving the hypothesis a place in the order of precedence. … But experience must be our chart in economical navigation. (CP 7:220–22)

While it may seem odd to consider probability of truth under the heading of "economy," one way of reading Peirce is as expressing an important insight about how background knowledge, as well as practical constraint, informs how science should be (and not, coincidentally, often is) practiced (for a discussion, see Kronz and McLaughlin 2005).

When scientists consider which hypothesis to adopt and test, they are not entirely free to consider any possible hypothesis. One only adopts a hypothesis that one could, conceivably, test, under whatever practical constraints are in play. Moreover, Peirce remarks that "instinctively" (i.e., at first) a hypothesis appears "simple"—apt, reasonable, or "of good sense" to us. The hypothesis may have instinctively appeared so, exactly because we must have reason to believe it probable. That is, we may have good inductive reasons for accepting it—it accords not only with our preconceived ideas, wherein these preconceived ideas are often founded on theoretical or empirical knowledge that forms the background against which new ideas are proposed or tested. However,

it also may just be the first best option in a series of choices that are driven, in part, by the cheapness and ease of testing. Abduction, on this account, involves a process, including stages of reflection both practical and evidential.

2.4. The Descriptive Problem

This point leads naturally to the descriptive problem: characterizing the role of abductive inference in scientific practice. Peirce was always cognizant that a formalization of scientific reasoning is an idealization. A formal representation can only apply to scientific practice if we treat such practice in a stepwise process in which each stage may be clearly demarcated. Peirce was skeptical that such a demarcation was possible. He was concerned with the practice of scientific reasoning. This is, at least in part, what led him to the view that induction and deduction are insufficient, by themselves, to characterize scientific practice. The logic of abduction is absolutely central, in his view, not only to the first stages of science but also to the possibility of any kind of knowledge whatsoever. He writes that "it is the only logical operation which introduces any new idea" (EP 2.216). Peirce argues that abductive inference played a role in the discovery of the kinetic theory of gases and Kepler's positing of an elliptical orbit of Mars. Without abduction, in his view, we would be at sea—infinite possible hypotheses would present themselves, with no plausible ranking. Abduction is that "sorting" process that leads us to a hypothesis worth exploring further.

I have characterized four puzzles Peirce bequeathed the literature on abduction: the formalization indeterminacy problem, the boundary problem, the justification problem, and the descriptive problem. With this characterization of Peirce's puzzles in place, the project is twofold: First, I will briefly elaborate on some central debates in the history of the literature on abduction. Second, I will argue that many of the debates about the scope and warrant of abductive inference since Peirce have hinged on prior commitments to independent questions, for example, concerning whether philosophers of science ought to be concerned with the psychology of discovery and how or whether empirical psychology bears on our understanding of scientific reasoning.

3. Discovery or Confirmation?

Peirce argues that in abductive reasoning we do not infer to, that is, accept as true, or even assign probabilities to a given hypothesis. Rather, he argued that in abduction, we are accepting a hypothesis "only problematically, that is to say, as meriting an inductive examination" (CP 2:786), or, as he says elsewhere, "It

236

is one act of inference to adopt a hypothesis on probation. Such an act may be called an abduction. It is an act of the same kind, when a hypothesis is merely suggested as possibly worth consideration. For even then some degree of favor is extended to it. But when, in consequence of having found that a good many predictions based on the hypothesis have been verified, a man begins to have positive belief in it, that is an act of inference of a totally different kind: it is an induction, or reasoning from a sample."[2] This passage suggests that Peirce would view abduction as distinct from most contemporary accounts of inference to the best explanation (IBE). For many, if not most, accounts of IBE take it to be thoroughly in the context of confirmation. Yet, according to the above, once we take a hypothesis to be not only worth exploring but well confirmed, we have moved from the context of abduction to induction.

Admittedly, Peirce struggled with whether these two forms of inference are genuinely distinct. He writes, "Induction is, plainly, a much stronger kind of inference than hypothesis; and this is the first reason for distinguishing between them. Hypotheses are sometimes regarded as provisional resorts, which in the progress of science are to be replaced by inductions. But this is a false view of the subject. Hypothetic reasoning infers very frequently a fact not capable of direct observation. … How is that hypothesis ever to be replaced by induction?" (EP 1:198). Here, Peirce is struggling with several independent features of inductive versus abductive inference; he seems to suggest at first that the distinction is a mere matter of strength of inference, but then he qualifies his statement, suggesting that the latter is distinctive in inferring from observable to unobservable or from instances to a rule versus from complex facts to an explanatory theory. His view here seems to be that the transition from accepting a hypothesis "on probation" to viewing it as "verified" is a gradual one.

Peirce often characterizes abduction as distinctive in that it results in a rather different course of action than mere induction, namely, the further investigation of a given hypothesis. Some take this to suggest that Peirce is advocating a kind of "practical reasoning," recommending a "course of action," as opposed to an epistemic state, as a result of abduction. Indeed, Hintikka (1998) argues that abductive inference is not inference at all; it is instead a kind of questioning strategy. And, Kapitan argues that "abductive-preference is a special type of practical reasoning, which culminates not in acceptance of H per se, but in an endorsement of a course of action" (1992, 14). In other words, on this line of interpretation, abduction is a strategic practice (involving merely plausible

2. Peirce's unpublished writings, Annotated Catalogue, ed. Richard Robin, Institute for Studies in Pragmaticism, MS 873.23, available in the microfilm edition of Peirce's papers at Harvard. Quoted with permission from the Harvard University Department of Philosophy.

belief) that results in a proposed course of action—investigating a hypothesis further. That is, on this reading, Peirce is not (as he avows) advancing a new form of inference but, rather, a form of life, a practice for science.

Kapitan follows his characterization of Peirce's "abduction preference" as "practical reasoning" with the following statement: "The proper conclusion might be that he lacked a clear and consistent picture of what is to be inferred in abduction, or, for that matter, a clear distinction between belief and other practical attitudes" (1992, 15). Peirce surely saw belief and practical attitudes as connected. One often has good reason to "hold" a hypothesis, at least tentatively, in the course of science, for the purposes of exploration, before one has reason to "believe" it. This seems to mark off abduction from IBE.

However, some have argued that abductive inference and IBE are one and the same. Niiniluoto, for instance, simply takes the two to be of a piece and dispenses with any controversy over this issue quite quickly. He cites Peirce's claims that "abduction consists in studying facts and devising a theory to explain them" (CP 5:145) and "abduction is the process of forming explanatory hypotheses" (CP 5:172) as evidence for this reading. Further, if one looks to Peirce's formalization of abduction ("The surprising fact, C, is observed. But, if H were true, C would be a matter of course, hence. There is reason to suspect that H is true") and supplies the following third premise, as a corrective—"H is more economical than the envisioned competitors," where economical has to do with not only cost or simplicity but consistency with prior (empirically grounded) belief—the argument sounds much more like an inference to the "best" explanation. The above passages certainly suggest that abductive inference is a form of IBE. And, it is true that Peirce stresses "explanationist" virtues like breadth and "incomplexity" when assessing hypotheses (CP 7:220–22).

Peirce surely believed that abduction is guided by explanatory virtues. However, if it results in a state of belief, this belief is not necessarily that the hypothesis in question is "probable," in the sense of "better confirmed"; rather, it is plausible in the sense of "worth investigating" or "provisional." Peirce is concerned to characterize the arrival at a theory worth exploring, in the context of discovery. However, he also realizes that in characterizing scientific reasoning, he is characterizing a process—something that transgresses the boundary between the context of discovery and the context of justification. Explanationist virtues arguably play an important role in both early stages of assessing competing hypotheses and later ones.

In contrast, IBE, at least on most contemporary accounts, is more concerned with the context of justification (see, e.g., Lipton 2004). That is, most philosophers of science today who are concerned with IBE are concerned with it as a rule of inference for confirmation. Thus, while abduction is like IBE

in that it involves determining whether and how the hypothesis would explain the phenomenon in question, it is unlike IBE in that it does not result in acceptance of a hypothesis as either true or (even) highly probable. It is simply worth investigating further. That abduction seems somehow lesser in aim than IBE leads naturally to the question of whether explanationist virtues are enough to justify belief. This latter concern has led to a great deal of the criticism of abduction in the philosophical literature.

4. Abduction's Warrant

Below, I will discuss two different views on abduction's warrant. Essentially, this involves pushing abduction over into either the context of justification and trying to spell out in what sense arguments of this sort are warranted or treating it as firmly in the context of discovery. Admittedly, these are extremes along a continuum; it may be that no one has occupied either extreme, but they illustrate two dominant tendencies in the literature on abduction and have resulted in very different assessments of its warrant.

One extreme "reductionist" view is that abduction is either incoherent or can be reduced to some other form of reasoning (e.g., induction, Bayesian updating, relative likelihoods). Abduction is simply inferring to the likeliest explanation, where this may be understood as that hypothesis most strongly confirmed by the total evidence, at the stage at which we are concerned with confirmation. This view has often (but need not necessarily) been defended alongside the view that we ought to jettison any hope for a logic of discovery. Rather, one ought to treat abductive inference as just another form of confirmatory inference—admittedly one that appears to deploy explanationist virtues. However, the extreme reductionist stance generally assumes that any "mysterious" explanatory properties of a theory can and should be reduced to or explained in terms of empirical adequacy. We may and should formalize abduction (sometimes understood as just IBE, by another name), using one of the formal models of confirmation, for example, Bayesian updating or likelihood ratios.

One of the most famous critics of abductive inference is Popper. Popper denied that there could be any logic of discovery: "The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it" (1934/1968, 31). Popper seems to be the most thoroughly reductionist of all; the forming of hypotheses is not subject to reason, on this view. However, Peirce, to some extent, anticipates this line of objection and attempts to refute it. He writes: "Think of what trillions of trillions of hypotheses might be made of which one only is true; and yet after

two or three—or at the very most a dozen guesses, the physicist hits pretty nearly on the correct hypothesis. By chance he would not have been likely to do so in the whole time that has elapsed since the earth was solidified" (CP 5:172). Peirce believes that were there no logic of discovery—that is, no methods for the arrival at hypotheses worth investigating further—scientists could never begin, much less finish, scientific inquiry. That is, if hypotheses were proposed entirely at random, one would have to wait until the end of the earth for science to even begin. Peirce concludes there must be some logic or, at least, heuristic for scientific discovery. Abduction is this method.

5. The Heuristic Stance

Until relatively recently, one of the few figures that took up Peirce's cause was Norwood R. Hanson. In *Patterns of Discovery* (1958/1972), Hanson echoes Peirce in the judgment that Kepler's discovery of the inverse-square law as one of the most important retroductive inferences in the history of science. Hanson goes to great pains to describe the particular details of Kepler's reasoning about Mars's orbit and how, at great length, he finally came to the elliptical figure. However, Hanson does not seem particularly interested in the formalization or the justificatory problem, aside from arguing that what he calls the "inductivist" and the HD account (Hanson uses "HD" to refer to a hypothetico-deductive account of science, wherein observations are deduced from general laws; 1958/1972, 71) both fail as an account of such processes as Kepler's or "how laws are come by in the first place." They fail as an account of "physical practice." The HD account is helpful when discussing "the argument of a finished research report," and the inductivist account is acceptable where laws are arrived at "by enumeration of particulars," but both fail to capture the "first tentative proposal of laws," the "reasons for proposing or trying out a hypothesis" (70–71).

In other words, Hanson thinks that both inductive and deductive forms of reasoning fail to capture something important about the practice of science. Yet, in claiming this he insists that he is not claiming that scientific discovery is not reason guided. Hanson explains that there is a process of reasoning leading to scientific theories—"retroduction"—reasoning from "explacanda to explacans" (1958/1972, 85, 71). Kepler's process of reasoning leading to the elliptical orbit of Mars, "what Peirce calls the finest retroduction ever made" (73), was a matter of considering possibilities first as mathematical idealizations, and then as real possibilities. Hanson describes Kepler's reasoning thus: "If we consider the orbit to be elliptical, then … let us suppose that our figure were a perfect ellipse. … Let us see what follows therefrom" (79). This mode of

reasoning seems to parallel Peirce's if-then formulation of abductive inference, from 1903.

However, Hanson's view was subject to a number of criticisms in the 1960s, '70s, and '80s. To claim that "retroduction" was reason guided, but not amenable to the formalizations of confirmation then on offer, made the problem of describing a logic of discovery, if such there was, very attractive. In the 1970s and '80s, the main representatives of this project were Gutting (1973, 1980), Thagard (1978), Nickles (1980a, 1980b), and Achinstein (1985). However, all of these authors more or less assimilated abduction and IBE, bringing their view more in line with the reductionist stance. Even Niiniluoto (1999), more recently, sees abduction as explanatory inference. However, as we have seen above, "acceptance" of explanations was not what Peirce intended by abduction, which was only tentative entertainment, relative to the alternatives.

6. Abduction and Confirmation

At first, Harman (1965) argued that all inductive inference is inference to the best of competing explanations. Even enumerative induction, in Harman's view, is a species of IBE. That is, when a person infers a generalization, his inference is good to the extent that the generalization offers a better explanation of the evidence than competing hypotheses do. Harman does not, at least at first, give a formal account of how we are to judge one hypothesis as providing a better explanation than another. Rather, he offers informal criteria, for example, a hypothesis should be simpler, less ad hoc, and so on. That is, at this stage, Harman seems very much in the antireductionist camp—appealing to explanationist virtues like simplicity without attempting to formalize them.

However, in 1967, Harman is far more precise. An explanation is better than competing explanations when the truth of the hypothesis in question would have made the outcome more likely—or, more formally, when the degree of confirmation (c) of some evidence (e(h)) entailed by the total evidence (e) would be higher if h were true than if not-h were true ($c((e(h) \text{ if sub } h), e) \gg c((e(h) \text{ if sub not-}h), e)$; Harman 1967, 411). Harman is using c here to refer to "degree of confirmation," and "e(h) if sub h" refers to the subjunctive conditional: e is true under the assumption of h. Essentially, Harman seems to hold that assessment of relative explanatory power is a matter of likelihood ratios.

Harman (1973) substitutes an "explanatory" for Goldman's "causal" theory of knowledge, again appealing to IBE as a way of addressing Gettier problems. However, both of these moves assimilate abductive inference to IBE, and IBE to relative likelihood, such that it becomes unclear whether abductive inference

is distinct from other forms of confirmation. Abduction is no longer in the context of "discovery" but is thoroughly in the context of "justification."

Similarly, van Fraassen (1980) places IBE in the context of justification and argues that it fails as a method of confirmation. For, if we use the Bayesian method of updating, and give an "additional boost" to hypotheses that are better explanations (above and beyond the probabilities of such hypotheses founded on evidence), then we will be subject to a Dutch book argument. Van Fraassen essentially takes IBE to be a rule of induction—and a rule for updating, prescribing how to update our opinions in view of evidence. It is Bayes's theorem, plus extra for explanatory considerations.

In contrast, Thagard argued for a variety of explanatory virtues as playing a role in theory choice: consilience, simplicity, and analogy. Thagard goes some way toward giving precise definitions of each of these notions. However, and this is very important, he insists that at least part of the process of inferring to the best theory involves pragmatic considerations. That is, for example, in the case of consilience, deciding which facts constitute a class or kind of fact depends importantly on historical context and background knowledge: "In inferring the best explanation, what matters is not the sheer number of facts explained, but the variety, and variety is not a notion for which we can expect a neat formal characterization" (Thagard 1978, 83). Likewise, with respect to simplicity, Thagard argues that it is a function of the size and the nature of the set of auxiliary hypotheses; how we count auxiliaries is again a matter for the messy historical details and background knowledge in any particular field. More recently, Thagard has developed a computational theory of explanatory coherence. On his view, an account of abduction should "[go] far beyond the account that takes abduction to be a form of inference: q, if p, then q; maybe p" (Thagard 2007, 228). He argues that there are neural mechanisms for abduction and that positive emotions are associated with the discovery of a hypothesis.

Thagard takes both discovery and justification of scientific theories to often be the result of abduction. Abductive inferences, thus, may be weak or strong; stronger such inferences involve multiple abductions, showing that a theory explains a variety of facts. The examples of Darwin and the wave theorists show that analogies figure in arguments concerning the best explanation. Because analogy is a factor in choosing the best explanation, there is no logic of discovery distinct from the logic of justification. Analogy may be used either to direct inquiry toward certain kinds of hypotheses or to support hypotheses already discovered (Thagard 1978, 90). This view is far closer to that of Peirce (see, e.g., EP 1:196–97, 1:300). However, Peirce does take pains to distinguish arguments from hypothesis from arguments from analogy (EP 1:34).

Lipton has argued (2004) that IBE is a distinctive form of nondeductive inference. On Lipton's view, there is an important distinction between likeliness and loveliness of a purported explanation. What makes an explanation lovely is the possession of such theoretical virtues as simplicity and unifying power. IBE and Bayes are consistent, not competing accounts of nondeductive inference. One might coordinate (but not reduce) theoretical virtues associated with the 'best explanation' with Bayesian priors and likelihoods—"loveliness" might be a guide to "likeliness." In other words, Lipton's is a nonreductive account.

What all these competing views have in common is an attempt to balance two goals that have, historically, been somewhat in tension in philosophy of science. On the one hand, one wishes to offer a systematic and normative account of how science "should" go, if one wishes to argue that scientists are empirically guided folk, with good reasons for their choices. On the other hand, philosophers from the '60s through the '80s were facing a series of challenges from historians and sociologists of science to the effect that scientific communities were guided by values and virtues that were not always reducible to, or explainable in terms of, empiricist virtues. That is, it appeared, at least according to the new historicism, that science does not always advance by a steady march of confirmation by more and better evidence. Arguably, the debate about abduction, when and whether it is warranted, and how was reflecting a larger debate in philosophy of science more generally, about how to acknowledge historical insight into the nature of scientific practice but maintain the "philosopher's" role of setting out normative standards for scientific reasoning.

7. Compatibilism: A Middle Ground between Reductionism and the Heuristic Stance?

An interesting recent trend in the literature on IBE attempts to reconcile these competing tendencies. What I will call the "compatibilist" trend treats the reductive and nonreductive views as compatible, rather than at odds (Okasha 2000; McGrew 2003; Lipton 2004). On this view, IBE is a sort of heuristic, helping us to meet Bayesian standards. That is, explanatory considerations are shorthand for what amount to appeals to priors.

However, Weisberg (2009) has argued that this solution is no solution at all, insofar as it holds IBE hostage to subjectivist Bayesianism, essentially "robbing IBE of much of its substance and interest" (125). In other words, according to Weisberg, those who take IBE to be reducible to Bayesian confirmation are, essentially, attempting to solve the formalization indeterminacy problem, perhaps at the expense of the descriptive problem. Similarly, Schupbach and

Sprenger (2011) argue that a probabilistic measure of explanatory power may be modeled formally. Essentially, theirs is a strategy of "saving" IBE without reducing it to Bayes's rule.

8. Conclusion

The history of the controversy over abduction often has to do with an exclusive focus on one of the several problems identified above, to the exclusion of others. The reductionists focus exclusively on what I have called Peirce's formalization indeterminacy and justification problems, pushing the problem of abduction (sometimes assimilated into IBE) into the problem of characterizing the context of justification, not the context of discovery. Neo-Hansonians attempt to save abduction as a process of discovery, sometimes at the expense of addressing any justificatory problems.

However, both extremes suffer. Reductionists are correct to worry about the problem of a warrant for abduction. That is, abduction is an incredibly weak form of argument that does not give us any definite probabilities or likelihoods. However, this would have been no surprise to Peirce. According to Peirce, abduction is at best useful in prescribing a course of action, namely, the tentative adoption of a belief, to guide further investigation.

Abduction does involve weighing hypotheses one against another, due to both evidential considerations and other more pragmatic considerations. So, Bayes's rule and likelihood ratios approximate, but do not capture in full, abductive inference's role in science, exactly because abduction is not concerned exclusively with testing or confirmation but with the extended strategic process of discovery and investigation of hypotheses worth exploring (Hintikka 1998; Paalova 2004, 2006). While I agree with reductionists that what count as good explanations or tentative hypotheses are importantly tied to background theory and evidence, it is impossible to find a principled way to map explanationist virtues onto priors and likelihoods. While reductionists may see this as a problem with explanationist virtues, an alternative view, which Peirce perhaps anticipated, is that this is a problem with the formalism, not the explanationist virtues. Formalizing what amounts to a strategic series of choices, often guided by pragmatic virtues like the ease or lesser expense of testing some hypotheses over others, seems to miss an important part of scientific practice.

While there are several adequate ways to formally represent abduction or IBE in terms of likelihoods and priors, the formalism largely underdetermines scientific practice. However, it does not, at least not obviously, follow that discovery is irrational or merely a random stroke of "insight." I venture to suggest that Peirce knew this well; the intertwined processes of discovery and

244

justification in science do not, at least not with any ease, admit of formalization. As Ben-Menahem argues, "when trying to lay down standards of explanatory merit, the philosopher often envisages a different type of standard than that which is commonly used in real day-to-day assessments of explanatory merit. Since he looks for structural features which are general enough to characterize all adequate explanations, he pays less attention to more detailed, nonstructural standards which each particular discipline develops and which, as a matter of course, vary with time and context" (1990, 325).

In sum, Peirce left a legacy of unanswered questions about abduction. Open questions about abduction include

Can or should we formalize abduction? If so, how?
Does abduction commit a simple error of reasoning, for example, affirming the consequent? Or, is abduction a form of inference that should be judged on different grounds from deduction?
Is abduction the same or different from induction? May one be reduced to the other? Which way does the reduction go?
Is abduction the same or different from IBE?
Is abduction a method of discovery, confirmation, or both?
Is abductive "discovery" formally or otherwise distinct from abductive "confirmation"?
Is there a continuous range of applications of abduction in science? If so, does the aim of abductive inference change by application or context—is it sometimes aimed at truth and sometimes at something rather different (e.g., a hypothesis worth investigating—for reasons including but not exclusively limited to probable truth)?
Is the narrowing of considerations to plausible hypotheses part of the process of abduction? Is it part of IBE? Or is it before both?
If IBE belongs in the realm of confirmation and abduction in the realm of discovery, when does abduction become IBE?

The above questions are open in part because we have not (once and for all) answered prior questions about the aim of philosophy of science. Is philosophy of science a matter of describing scientific practice or prescribing sound forms of inference or both? Can or should philosophy of science investigate scientific discovery or only contexts of justification; should or can they be treated independently? Should philosophy of science (like ethical theory of late) be instructed or constrained by empirical facts about how we do reason? If we find that we commonly use a form of inference as a rule of thumb that turns out to violate what we know or believe to be the canons of "good reasoning," should

we rethink our canons, our rationality, or both? The philosophical conversation about abduction has occasionally run into irresolvable deadlocks exactly because the authors of opposing views answer these questions differently. All these issues arise in the context of discussions of abduction exactly because Peirce left the door open to options that many today treat as closed.

REFERENCES

Achinstein, Peter. 1970. "Inference to Scientific Laws." In *Minnesota Studies in the Philosophy of Science*, vol. 5, ed. R. H. Stuwer, 87–111. Minneapolis: University of Minnesota Press.

———. 1980. "Discovery and Rule Books." In *Scientific Discovery, Logic, and Rationality*, ed. Thomas Nickles, 117–37. Dordrecht: Reidel.

———. 1985. "The Method of Hypothesis: What Is It Supposed to Do, and Can It Do It?" In *Observation, Experiment, and Hypothesis in Modern Physical Science*, ed. P. Achinstein and O. Hannaway, 127–45. Cambridge, MA: MIT Press.

———. 1987. "Scientific Discovery and Maxwell's Kinetic Theory." *Philosophy of Science* 54:409–34.

Anderson, Douglas. 1986. "The Evolution of Peirce's Concept of Abduction." *Transactions of the Charles Sanders Peirce Society* 22:145–64.

Ben-Menahem, Yemima. 1990. "The Inference to the Best Explanation." *Erkenntnis* 33:319–44.

Burton, Robert. 2008. *On Being Certain: Believing You Are Right Even When You're Not*. London: St. Martin's.

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

Frankfurt, Harry. 1958a. "Peirce's Account of Inquiry." *Journal of Philosophy* 55:588–92.

———. 1958b. "Peirce's Notion of Abduction." *Journal of Philosophy* 55:593–97.

Gopnik, Alison, and Laura Schulz. 2007. *Causal Learning: Psychology, Philosophy and Computation*. Oxford: Oxford University Press.

Gutting, Gary. 1973. "A Defense of a Logic of Discovery." *Philosophical Forum* 4:384–405.

———. 1980. "Science as Discovery." *Revue Internationale de Philosophie*, nos. 131–32: 26–48.

Hahn, U., and M. Oaksford. 2007. "The Rationality of Informal Argumentation: A Bayesian Approach to Reasoning Fallacies." *Psychological Review* 114:704–32.

Hanson, Norwood Russell. 1958/1972. *Patterns of Discovery*. Cambridge: Cambridge University Press.

Harman, Gilbert. 1965. "The Inference to the Best Explanation." *Philosophical Review* 74:88–95.

———. 1967. "Detachment, Probability, and Maximum Likelihood." *Nous* 1:401–11.

———. 1968. "Knowledge, Inference, and Explanation." *American Philosophical Quarterly* 5:164–73.

———. 1973. *Thought*. Princeton, NJ: Princeton University Press.

Hintikka, Jaako. 1998. "What Is Abduction? The Fundamental Problem of Contemporary Epistemology." *Transactions of the Charles S. Peirce Society* 34 (3): 503–33.

Kapitan, Tomis. 1992. "Peirce and the Autonomy of Abductive Reasoning." *Erkenntnis* 37:1–26.

———. 1997. "Peirce and the Structure of Abductive Inference." In *Studies in the Logic of Charles Sanders Peirce*, ed. Nathan Houser, Don D. Roberts, and James Van Evra. Bloomington: Indiana University Press.

Kraus, Manfred. 2003. "Charles Peirce's Theory of Abduction and the Aristotelian Enthememe from Signs." In *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*, ed. Frans H. van Eemeren et al. Amsterdam: Sic Sat.

Kronz, Fred, and Amy McLaughlin. 2005. "The Fast Track to Confirmation: Achinstein and Peirce on Confirmation." In *Scientific Evidence: Philosophical Theories and Applications*, ed. Peter Achinstein, 69–84. Baltimore: Johns Hopkins University Press.

Lipton, Peter. 2004. *Inference to the Best Explanation*. 2nd ed. New York: Routledge.

McGrew, Timothy. 2003. "Confirmation, Heuristics, and Explanatory Reasoning." *British Journal for Philosophy of Science* 54:553–67.

Nickles, Thomas. 1980a. "Scientific Discovery and the Future of Philosophy of Science." In *Scientific Discovery, Logic, and Rationality*, ed. Thomas Nickles, 1–59. Dordrecht: Reidel.

———, ed. 1980b. *Scientific Discovery: Case Studies*. Dordrecht: Reidel.

Niiniluoto, Illka. 1999. "Defending Abduction." *Philosophy of Science* 66:S436–S451.

Okasha, Samir. 2000. "Van Fraassen's Critique of Inference to the Best Explanation." *Studies in History and Philosophy of Science* A 31 (4): 691–710.

Paalova, Saami. 2004. "Abduction as a Logic and Methodology of Discoveries: The Importance of Strategies." *Foundations of Science* 9 (3): 267–83.

———. 2006. "Hansonian and Harmanian Methods as Models of Discovery." *International Studies in the Philosophy of Science* 20:93–108.

Peirce, Charles. 1931–58. *Collected Papers of Charles Sanders Peirce*. Vols. 1–6, ed. C. Hartshorne and P. Weiss; vols. 7–8, ed. A. W. Burks. Cambridge, MA: Harvard University Press.

———. 1982–2010. *Writings of Charles S. Peirce: A Chronological Edition*. Vols. 1–8, ed. the Peirce Edition Project. Bloomington: Indiana University Press.

———. 1992–98. *The Essential Peirce: Selected Philosophical Writings*. Vol. 2, 1893–1913, ed. the Peirce Edition Project. Bloomington: Indiana University Press.

Popper, Karl. 1934/1968. *The Logic of Scientific Discovery*. New York: Harper & Row.

Salmon, W. 2001. "Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation." In *Explanation: Theoretical Approaches and Applications*, ed. Giora Hon and Sam S. Rakover, 61-91. Dordrecht: Kluwer.

Schupbach, Jonah N., and Jan Sprenger. 2011. "The Logic of Explanatory Power." *Philosophy of Science* 78 (1): 105–27.

Sober, Elliott. 1994. "Contrastive Empiricism." In *From a Biological Point of View*, 114–36. Cambridge: Cambridge University Press.

Thagard, Paul. 1978. "The Best Explanation: Criteria for Theory Choice." *Journal of Philosophy* 75:76–92.

———. 2007. "Abduction: From Philosophical Analysis to Neural Mechanisms." In *Inductive Reasoning: Experimental, Developmental and Computational Approaches*, ed. A. Feeney and E. Heit, 226–47. Cambridge: Cambridge University Press.

van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Weisberg, Jonathan. 2009. "Locating IBE in the Bayesian Framework." *Synthese* 167 (1): 125–43.