20

Abduction and Inference to the Best Explanation

John Woods

> … Peirce is a major star in the firmament of philosophy. By thrusting the notion of abduction to the forefront of philosophers' consciousness he created a problem which—I will argue—is the central one in contemporary epistemology.
>
> ❖ Jaakko Hintikka (2007)

1. Critical thinking

Arising from the Greek word *"kritikos"* for critic, critical thinking is said to be the intellectual capacity and wherewithal for judgment and discernment. According to the United States National Council for Excellence in Critical Thinking, critical thinking is the intellectually, analyzing, synthesizing, or evaluating information gathered from or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action.[1]

Some writers tend to view it not as a first-order mode of reasoning or family of such modes, but rather as more of a meta-reasoning mode of reflection, designed to maximize the rationality of human thinkers. In the opinion of Carmichael Kirby, the human thinker doesn't employ critical thinking to solve problems. Critical thinking is used to improve the process of thinking (Kirby 1997). From that point of view, a critical thinking textbook could be seen as a manual of self-help for the ratiocinatively challenged.

Strictly at odds with the meta-reasoning approach is the opinion of Kerry Walters, who proposes that rationality must be more than the ability to operate the "calculus of justification" and must include "cognitive acts such as imagination, conceptual creativity, intuition and insight" (Walters 1994, p. 63). These capacities focus more on discovery than on linear, rules-based procedures for problem solving. The linear and non-sequential mind in a recipro-

---

[1] "Critical Thinking Index Page" at http://www.criticalthinking.org//

cally nourishing partnership is what a human being's full rationality requires. Walters's point enjoys a fairly substantial empirical backing from various of the sciences of cognition. It raises a methodological question for philosophers: Is there an established approach to epistemology that gives this feature of "full rationality" the heed that's due it? For the most part by far, CT theories operate on background "command and control" epistemologies. These are epistemologies in which the executive centre of reasoning and decision is the human intellect, operating of its own volition to run the engines of inference and choice. Polar opposites of command and control epistemologies are "causal response" theories in which the cognitive devices that do most of the work of inference, belief-change and decision, operate subconsciously—out of sight of the mind's eye, beyond the reach of the heart's command, and unengageable by tongue or pen (or keystroke). Given that critical thinking is an indispensable facilitator of knowledge, CT theories have a stake in those epistemological issues.[2] If, as I myself believe, Hintikka is right about the centrality of the logic of abduction to contemporary epistemology, CT theorists would also have a stake in an appropriately epistemologized account of abduction.

My remit here is to reflect upon the present state of theories of abductive inference (A) and inference to the best explanation (IBE). Since this chapter is part of a volume on tools of critical thinking, it behooves me to say a little something of the connection, or want of it, between the two research literatures. The abduction and inference-to-the-best-explanation literatures are large and wide-ranging. The critical thinking literature(s) range even more widely. But, for all their respective robustness, there is little in the way of productive intersections of them. Not only don't these literatures talk to one another (much), they don't even talk all that much to themselves. However, on the whole the A&IBE literatures are the more interactive and unified of the two. One of the reasons that the CT literatures are less unified is that the disciplines in which they are sited, don't normally talk to each other. The three dominant disciplinary sectors of the present-day CT literature are those that cater capaciously for informal logic, argument, and speech communication.[3] In Chapter 1 in this volume you will find a summary of where the CT literatures are to be found. All that there is space for here is to make some mention

---

[2] For a fuller discussion see (Woods 2017).

[3] Two somewhat more recent developments should be noted, both offered as contributions to the theory of argument. One is the emergence of highly complex mathematical and computational models of attack-and-defend arguments. See here the contributions of Barringer *et al.* (2002). The other is an even more expansive move to legal reasoning. See, for example, Douglas Walton (2008) and John Woods (2015).

of where contributors to the CT literatures hook up with A&IBE agendas. In the section below, I'll list some of the leading A&IBE works, indicating those scant few by authors of some prominence in the CT sector.

On the other hand, those parts of the CT literature in which we might expect to find some mention of A&IBE have so far left few footprints on the A&IBE literature and none that is either weighty or enduring. So the point is made: These are literatures that have yet to talk to each other about these matters of common interest. This helps in framing a course of action for what lies ahead in this note. A good part of what I aim for here is help to open an A&IBE backchannel to CT.

2. Peirce on abduction

We owe the modern notion of abduction to C.S. Peirce (1839-1914), although there are undeveloped intimations of it in 4th century B.C. in Aristotle's concept of *apagogē*. These need not detain us here (see Aristotle 1985). Peirce famously sees abduction as inference in the form,

(1) The surprising fact *C* is observed.
(2) But if *A* were true, *C* would be a matter of course.
(3) Hence there is reason to suspect that *A* is true. (*CP* 5.189)[4]

Peirce's schema covers not all of what an abductive inference is, but only what I'll call the "inference stage". It is preceded by the "selection stage". At the inference stage a claim is inferred about a hypothesis, namely, that if premises (1) and (2) are true, it follows that there is reason to suspect that a given hypothesis *A* is true. At the selection stage—not schematized here—Peirce reflects on how hypotheses are arrived at in the first place. The logistics of neither stage are all that easy to discern, but the selection stage poses the rougher challenge by far. Here is more of what Peirce had to say about them.

(i)   Abduction is triggered by surprise. (*CP* 5.189)
(ii)   Abduction is a form of guessing, underwritten innately by instinct. (*Reasoning and the Logic of Things,* p. 128.[5] *CP* 5.171, *CP* 7.220)
(iii) Abduction provides no grounds for believing the abduced proposition to be true. (*RLT,* p. 178)

---

[4] C. S. Peirce, *Collected Works*, Cambridge, MA: Harvard University Press, 1931-1958. A series of volumes first appearing in 1931. "*CP* 5.189" denotes p. 189 of volume five. Line numbers are mine.

[5] C. S. Peirce, *Reasoning and the Logic of Things: The Cambridge Conference Lectures of 1898,* K. L. Kettner, editor, Cambridge, MA: Harvard University Press, 1992. Hereafter "*RLT*".

(iv) Rather than believing them, the proper thing to do with abduced hypotheses is send them off to experimental trial. (*CP* 5.599, *CP* 6.469-6.473, *CP* 7.202-219)

(v) The connection between the abduced hypothesis *A* and the observed fact *C* is subjunctive. (*CP* 5.189)

(vi) The conclusion licensed by the abduction is not proposition *A*, but rather that *A*'s truth is something there is reason to suspect.

(vii) The "hence" of the schema's conclusion is not characterized at *CP* 5.189, but it is clear from context that it denotes a nonmonotonic relation.

The present-day literature reveals different streams of thought about bringing the Peircean paradigm into a more unified and systematic theory. There isn't space in this note to adjudicate these differences. Interested readers could consult to advantage the somewhat selective works noted in the Appendix, arranged there in chronological order.

Abductions of Peirce's sort are entirely commonplace, and the fact that human beings are fairly good at drawing them makes a large contribution to the success of *homo sapiens* in negotiating the shoals of life's dangerous waters. Imagine Fred, a middle aged man, married to Kate, an accountant, whose three grown children live away, two married and with busy careers, and a third who is a student in residence at an out-of-province university (the family lives in Canada). One December day Fred returns from his Court Clerk's job at his usual time of five, parks in the garage at the back. As he starts walking towards the rear of the house, he sees that the back door is wide-open. The downstairs lights are on, triggered to do so automatically around 5:00 p. m. This is the door from which Fred leaves the house every morning on the dot of 6:30 for a pre-work visit to the gym. He has never before failed to close it properly. Kate is driven to work by a neighbour, and always leaves through the front door. Her own unvarying after-work routine takes her from office to gym, and from there to the market where she does her daily shopping, finally arriving home by bus at 6:00. It is said that the philosopher Kant was so regular in his walks that the citizens of Königsberg could set their watches by when he started or ended. Fred and Kate were like that too.

In Peircean terms, the open door was the surprising fact *C*. What Peirce means by a "surprising" one is a fact that is somehow off-course, irregular, or aberrant. The open door was not how things were supposed to be at Fred and Kate's house. The open door also posed an *ignorance-problem* for Fred. He wanted to know why it was open, and yet nothing he presently knew answered his question. Fred's ignorance wasn't something to make light of. There were innumerable things Fred didn't know then and there, such as the

size of Montenegro's population, for example. Of course, had he wanted to know it, he could have pulled out his iPhone and found out in a flash. The open door was different. It was a cognitive irritant, and was not something Fred could or should simply have put aside. And no answer awaited him on his iPhone.

Given this cognitive irritant (not necessarily something psychologically annoying), Fred wanted to find a hypothesis *A* with respect to which, were it true, the open door would have been a matter of course. Suppose that *A* were the hypothesis that Fred had simply forgotten to shut the door securely. If that had been what actually happened, then with no one left in the house to spot it open, it would be entirely a matter of course that Fred would be the first to discover it later on. In Fred's household, a closed door at 5:00 p.m. is a reliable but not strictly universal generalization, certainly not universal in the logician's sense. Colloquially speaking, it was "never" open at 5:00 p.m. But while a perfectly accurate one so far, the regularity is defeasible. It admits of possible exceptions. What Fred wanted to know in this case was what caused the exception to this regularity. Not knowing of any, it was time for him to use his head. Was there, he wondered, a *hypothetical* fact, which, were it actually true, would have provided the exception that induced the open-door surprise? In so asking, Fred was looking for a Peircean hypothesis *A.*

It is here that a second difficulty arises. There are more possibilities than one to consider. It is possible that Kate had come home early and was in process of hanging some laundry on the outdoor clothesline. But there were two things wrong with this. In that same sense of "never" in which the back door is never left open, it is never the case that Kate comes home before Fred, and, Kate might for the first time have come home sick, but wouldn't that pretty well rule out the late-laundry hypothesis?

Perhaps Pete had come home from university for an unannounced weekend visit. But Pete was in the middle of final exams in a tough computer engineering programme. Besides, he'd never done that before, and would always have called ahead were he have decided to do so. Kate and Fred are entirely self-sufficient by choice. They share all household duties, including spring cleaning and all the rest. They do all their own lawn mowing, gardening and snow removal. No one has a key to the house beyond the five members of the family. Nice Mr. Chapman across the street once had one for emergency purposes. Childless and long retired, Mr. Chapman had recently died, and Mrs. Chapman shortly after had returned to her native England to live with her sister in Cambridge. The key had never surfaced, presumably now an inadvertent resident of that pleasant but chilly university town.

Fred and Kate's house hadn't been equipped with a security alarm. All three external doors were of solid construction set in sturdy frames, and pro-

vided with superior locks. The house lies in a quiet neighbourhood, and had never been broken into before. Nor had there been any evidence of any earlier attempt to do so. As he made his way towards the house, four different versions of line (2) of Peirce's schema might have somehow or other entered Fred's mind.

(2^1) If it were true that Fred had forgotten to shut the door this morning, then its being open now would have been a matter of course.[6]

(2^2) If it were true that Kate had come home early, then its being open now would have been a matter of course.

(2^3) If it were true that Pete had paid a surprise visit, then its being open now would have been a matter of course. (Pete is notorious for not shutting doors.)

(2^4) If it were true that the house has been burgled, then its being open now would have been a matter of course. (Flight, not door-shutting, is the first priority for house-breakers.)

The known facts at the time countervailed against (2^1) (2^2) and (2^3) rather more than against (2^4). Fred "never" leaves the door open. Kate "never" comes home early. Pete "never" pays surprise visits. Even though the house had never been burgled before, houses rarely, if ever, are burgled in Fred and Kate's neighbourhood. Just as, in one good sense of the word, it is a *norm* that Fred closes the door properly, a norm that Kate gets home after Fred does, that Pete stays put on campus at weekends. Here too there is a norm about house burglary. It normally happens that when a house is burgled in a neighbourhood like theirs, it will have been for the first time. Note, however, that the first three norms carry more predictive weight than the fact that Fred and Kate's house had never before been burgled, which carries little, if any. Still, each of these four is an antecedent of a subjunctive conditional statement, which *if true* would trigger the Peircean conclusion,

(3*) Hence there is reason to suspect that the antecedent of (2^1) [2^2, 2^3, 2^4] is true.

It is important to emphasize that a reason to *suspect* that a hypothesis is true is certainly not evidence that it *is* true, and not reason to *believe* that it's true either. Peirce is insistent in saying that a successful abduction provides no grounds for believing the abduced hypothesis (Peirce, *RTL*, p. 178). From which we may derive two further observations about the situation Fred was in.

---

[6] "Had it been true that X " is more idiomatic English than "If it were true that X had ". But because the latter more closely fits the wording of Peirce's schema, I'll stick with it.

The productivity of ignorance-preserving inference: Even if at least one of these subjunctive conditionals were true, the hypothesis thereby induced would not have solved Fred's ignorance-problem. But it might have provided Fred with a reasoned but non-deterministic basis for action.

Here is the action that Fred actually took. He pulled out his phone and called 911. He called Kate on her cellphone, advising her that, upon arrival, she shouldn't enter the house until the police had checked things out. He followed that advice himself. As it turned out, the house had indeed been burgled. The thieves had used a glass cutter to remove a small basement window.

We now have at hand a nice characterization of the good that's in a good abductive inference. It also helps distinguishing successful abductive inference from valid deductive inference and strong inductive inference. While

- a valid deduction is *truth-preserving,* and
- a strong induction is *likelihood-enhancing,*
- a successful abduction is *ignorance-preserving yet action-motivating.*

Suppose, however, that after careful examination of the scene, the police had found no evidence of a break-in and Kate and Fred had found nothing missing or damaged. At that point, they both might reasonably have defaulted to the hypothesis of the antecedent of subjunctive conditional (2^1) and drawn the conclusion:

(3*) Hence there is reason to suspect that Fred forgot to shut door.

It is interesting that in the discussion surrounding the presentation of the Peircean schema the key terms "surprise", "matter of course" and "suspect" are neither defined nor explicated. Even so, the wording of the schema provides some contextual clues. As I've already suggested, a surprising event is something off-course or irregular, something that deviates from the normal workings of things. In *CP* 1.139, Peirce says "Now nothing justifies a retroductive [= abductive] inference except its affording an explanation of the facts." He doesn't say what he means by "explanation", but it is easy to see it working in the schema as an alternative wording of "would make them a matter of course." I'll come back to this later when we discuss *inference to the best explanation* in Section 4.

I think that it may now be said that Fred's reaction to the open door hits all the buttons of what Peirce has told us so far about abduction. Perhaps the most striking feature of the case of the open door is how long it's taken to recount it. In presenting this sort of everyday situation to a class on critical thinking, it wouldn't be at all unusual for students to leap to the punch-line well before the case has been properly laid out. This is bears on what Fred

actually did. When he first spotted the open door, it was 5:04. When he started calling 911 it was 5:05. It took longer to answer the 911 operator's questions, hence longer to complete the call than to arrive at the point of making it. The situation that Fred was in is one of a kind that the students could easily see themselves being in. In each case, Fred's and their own, they moved like lightening. Yet every word laid out in Section 1 faithfully records most of what must have somehow or another been going through Fred's mind, and some likeness of what was going through the minds of the students to whom the Fred case is being recounted. Here are some highlights.

- There is a large presence of localized background information made up of numbers of defeasible household generalizations, supplemented by facts particular to those householders—e.g., that Fred knows what things are like in his city and in neighbourhoods like his.
- There is also the presence of a more global background information, having to do with Fred's largely implicit appreciation of the generalizations arising from his life-experiences to date and those of his ilk—e.g., open doors in December in northern Alberta, Canada make for cold houses.
- Also present were well-functioning filtration-mechanisms, keeping out of play myriad facts of Fred's explicit and/or implicit awareness that have no relevant bearing on the open-door problem—e.g., that Pete was born in August.
- There was also the fact that almost none of this relevant background information was processed consciously or formulated linguistically. From which we are invited to draw a further inference, perhaps the most important one so

The subconscious abduction thesis: Even though it is briefly and simply laid out, and Fred's actions comported with it, what he did not do was consciously instantiate the Peircean schema. There wouldn't have been time to call it to mind, and then give linguistic effect to the required substitutions for its schematic letters, in the interval between 5:04 and 5:05.[7]

Once again, this calls to mind the question of whether there is an epistemology that adequately takes this thesis into account.

This last point calls for a bit more consideration. Although what happened that day at Fred and Kate's place was a singular event, it was but one

---

[7] The contrast between consciousness and unconsciousness is certainly a graded one. It used to be thought (and still is in many circles) that unresponsiveness is a secure marker of unconsciousness. Recent work with a functional magnetic resonance imaging (*fMRI)* scanner casts doubt on this. See Adrian Owen (2017).

instance of a more commonly experienced type. The way in which Fred's mind worked on that occasion wasn't peculiar to reactions to Peircean surprises. It is typical of human responses to the information imparted by the varying elements of life's passing scene. Imagine a student in a critical thinking class, straining to take good notes at a lecture about Fred's case or anything else covered by the course. The student couldn't do it without a good aural understanding of the professor's language, a decent command of its syntactic structure and lexical orthography, as well as a timely appreciation of the necessity to get it all down while fresh. When we consider the magnitude of what the note-taking student must know, we should be struck by how little of it makes footfall in the student's conscious mind. For all its indispensability in getting the note-taking job done, by far its contribution was transacted in his cognitively productive subconscious. Nor is this peculiar to the cognitive mechanics of note-taking in a critical thinking class. It is, in Peirce's words, a matter of course for how human beings subconsciously process the information that conduces to their quotidian knowledge, calling to mind the importance of the non-sequential mind in Walters' approach to a well-nourished human rationality.[8] This, we may now think, has crucial implications for how theories of knowledge and logics of reasoning operate.

3. Big heavenly bodies

For all its commonplace appearance, the open-door abduction is a made-up example of one. One of the attractions of made-up examples is that they expose themselves to tendentious employment. You can make a made-up example do what you want, even to fit the provisions of an awaiting theory. Better, we might rightly think, to turn our minds to some real-life ones, beginning with the one that launched planet Neptune. We begin this story with Uranus, whose orbital irregularities were noticed in the 1840s to discomport with Newton's laws of motion and gravity. It was widely believed that these irregularities would be a matter of Newtonian course were there an as yet unknown planet whose gravitational force was disturbing Uranus orbit around the Sun. In 1845 Urbain Le Verrier (1811-1877) in Paris and John Couch Adams (1819-1892) in Cambridge, operating independently, made calculations to help account for the nature of this planet and to locate its position. Both Le Verrier and Adams drew the same abductive inference. The sought-for planet was Neptune. The postulation of Neptune spurred a flurry of follow-up testing, and seventeen days after the Royal Society had

---

[8] For a fuller discussion, see (Woods 2013 or Woods 2014). See also (Woods 2012 and 2017); this latter pair of papers update the account of (Gabbay and Woods 2005).

bestowed the Copley medal on Le Verrier, William Lassell (1799-1880) had tracked down Neptune's moon Triton. The Berlin Observatory made telescopic observations confirming Neptune's existence. It was a high point in 19th century astrophysics, and a stunning re-confirmation of Newton's laws.

We see in Neptune's discovery every box of the Peircean schema ticked off to rich astrophysical effect. The inference of Neptune was abductively impeccable, and its dispatch to the tribunal of observational determination yielded up a verdict that transformed a reason to suspect the abduced hypothesis to be true to a reason for asserting it to be so. Abductions so perfect are not, as Peirce makes clear, guaranteed such happy outcomes. We see this in the abduction of Vulcan.

It all began with the perihelion procession of Mercury's orbit whose classically predicted value differed from the observed one by 43 arcseconds per century. A small perturbation to be sure, but one that discomported with the Newtonian laws. In 1859, using the same techniques that had done so well in restoring Uranus to Newtonian fidelity, Le Verrier's team abduced the planet Vulcan, placing it between Mercury and close to the Sun. In the forty-five year aftermath of the abduction, there was no settled observational confirmation of Vulcan. In 1915, Einstein (1879-1995) published the general theory of relativity, which put Newton's laws into honourable retirement and, in so doing, revealed Mercury's orbital perturbations to be a relativity matter of course. To this day, there are astronomers who think that what was taken for a planet in 1859, might now be explained by other inter-Mercurial objects, either previously unknown comets or asteroids.

There are two points of importance to take note of here. One is that the abduction of Vulcan was as good as the abduction of Neptune. The other is that all abductions, Fred's or Le Verrier's or anyone else's, are transacted in contexts of background information, none of whose particulars functions as a voiced premiss. What happened between Neptune and Vulcan was a background shift from Newton to Einstein.

Another of our everyday defeasible generalizations is that everyone knows that water is H_2O. However, many fewer know what "hydrogen" and "oxygen" mean here and, still less, the structure denoted by the combined chemical notation. Vastly more people know that Einstein's general theory of relativity ended the reign of Newton's theory of gravity, than understand the new laws. It sounds paradoxical, of course. How could people know so much without knowing what they are talking about? An answer proposed in chapter 9 of *Errors of Reasoning* derives from the mediaeval philosopher-Saint Anselm (1033/4-1109), who famously avowed *credo ut intelligam*—"I accept

in order that I might come to understand."[9] The opposite view is the reverse avowal, "*intelligo ut credam*"—"I understand so that belief might be made possible." A question for epistemology in general is which, if either, of these Latin tags to favour.

One way out of the apparent paradox is that when everyone knows that water is H_2O, what they actually know is that "Water is H_2O" is a sentence that expresses a true proposition. Ditto for the likes of "Everyone knows that $e = mc^2$", and so on. On this characterization, it is perfectly consistent for someone to say that he really does know that *p* even if he doesn't know what he's talking about. This is as true of the man who first asserted the action at distance theorem, while acknowledging that he didn't understand it as it is true of the man across the street who says that $e = mc^2$, while acknowledging that he doesn't understand it. What Newton was saying, in effect, was "I don't understand the action at a distance theorem and yet there is reason to suspect that it expresses a true proposition." This, I think, is a good enough idea to have a name. We could call it "the Anselmian suspicion precedes understanding thesis", or for short "the suspicion of truth thesis". It would appear that the suspicion of truth thesis might have been purpose-built for a child's learning of the mother tongue. How likely that a little tyke would ever understand the utterances of parents and family without some antecedent primitive confidence of their reliability, buttressed by an also primitive testing of the abduction that generated the suspicion of the truth in the first place?

4. Explanation

In what must be considered one of the foundational papers, Gilbert Harman (1965, see pp. 88-95) picks up on an idea that we also see in Peirce's *CP* 1.139 (with apologies for repetitiveness):

> Now nothing justifies a retroductive [= abductive] inference except its affording an explanation of the facts.

Late in Section 2, I advanced the suggestion that what Peirce means by "explains" is wholly contained in premiss (2) of the Peircean schema. If so, a hypothesis *A* which, if true, would explain the surprising fact that *C*, would be the very one that makes *C* a matter of course. Peirce was never drawn to the idea that the inference to the best explanation is the keystone of scientific method. The very suggestion of it would have been dismissed upon arrival. The doctrine that suspicion of truth can both precede and facilitate subsequent

---

[9] Anselm, St., *Proslogion,* 1. This echoes remarks of St. Augustine of Hippo (354-430) in his *Tractates on the Gospel of John* 29.6: "*Crede*ut intelligas". , intelligas". *ut*ntelligas". *intelligas*

understanding is amply and convincingly instantiated in history. Neverthe-
less, Harman (1965) writes:

> 'The inference to the best explanation' corresponds approximately to what others
> have called 'abduction', 'the method of hypothesis', 'hypothetic inference',
> 'the method of elimination', 'eliminative induction' and 'theoretical inference'.
> (88-91)

Peter Lipton (1991) writes in like vein:

> According to Inference to the Best Explanation, our inferential practices are gov-
> erned by explanatory considerations. Given our data and our background beliefs,
> we infer what would, if true, provide the best of the competing explanations we
> can generate of those data (so long as the best is good enough for us to make any
> inference at all).

He goes on to say, quite rightly it seems to me, that

> Inference to the Best Explanation has become extremely popular in philosophi-
> cal circles, discussed by many and endorsed without discussion by many more .
> Yet it still remains much more of a slogan, than an articulated account of induc-
> tion. (*Ibid.*, p. 58)[10]

   We are now at the point of distinguishing what at present are the two most
discussed schematic updates of Peirce's original one. Like his, each of these
schematizes stage two—the inference stage—only. Let E be a Peircean sur-
prising event for some agent, K be his knowledge-base at the time, H be a
hypothesis, K(H) be a hypothetical knowledge-base got by addition of H to
K, and let ↬ be an undefined conditionality relation. Then, according to the
AKM schema, abductive inference has the following structure:[11]

---

[10] The same view persists in the 2004 edition of the 1991 book, *Inference to the Best Explanation.*

[11] The label "AKM" was proposed by Gabbay and me in our *The Reach of Abduction* 2005 in honor of some of those whose work is shaped by it, to wit: Aliseda 1997 and later 2006, Robert Kowalski, *Logic for Problem Solving,* New York: Else-vier, 1979, A. Kakas, R. A. Kowalski and F. Toni, "Abduction and abductive logic programming", *Journal of Logic and Computation,* 2 (1995), 719-770, Theo A. F. Kuipers, "Abduction aiming at empirical progress of even truth approximation leading to a challenge for computational modelling", *Foundations of Science,* 4 (1999), 307-323, Magnani 2001 and later 2009, and Joke Meheus, Liza Verho-even, Maarten Van Dyck and Dagmar Profijn, "Ampliative adaptive logics and the foundation of logic-based approaches to abduction", in Lorenzo Magnani, Nancy J. Neressian and Claudio Pizzi, editors, *Logical and Computational Aspects of Model-Based Reasoning,* pages 39-71, Dordrecht: Kluwer, 2002. Mind you, the

1. E
2. ~(K ↬ E)
3. ~(H ↬ E)
4. K(H) is consistent
5. K(H) is minimal
6. K(H) ↬ E
7. Therefore, H.[12]

A sizable percentage of the AKM community leans towards an explanationist interpretation of "↬" and a truth-preserving interpretation of "therefore" although variations have also been proposed. In some accounts, "↬" sometimes means "possibly explains", and "therefore" denotes high inductive likelihood, as reckoned up in the probability calculus. Whether "↬" means "explains" or "could explain", and "therefore" denotes truth-preservation or the enhancement of likelihood, there remains a question, any answer to which could be a hard sell. The question is, "What is an explanation?"

Hintikka writes in *Socratic Epistemology* that the idea that abduction is inference to the best explanation, while it "has a great deal of initial plausibility is seriously simplified at best." (p. 40) He continues,

> Part of the difficulty can be seen by asking, first what explanation is or, perhaps more pertinently, what explaining is. Most people who speak of 'inferences to the best explanation' seem to know what explanation is. In reality, the nature of explanation is scarcely any clearer than the nature of abduction. (*Ibid*., pp. 40-41.)

> In one sense—the hypothetico-deductive sense—"explaining an explanandum E is to derive it from [a] background theory T plus a number of contingent truths that are relative to E and that have to be found in order for an attempt to explain E to succeed" (*ibid*., p. 41).

If we accept this, it is easy to see that abduction is intrinsically *non*-explanationist. Let Hintikka's background theory T play the role of the AKM-schema's knowledge-base K and let Hintikka's derivation be the AKM "↬". On Hintikka's approach, the equivalents of H are contingent truths. But in

choice of the letters, "A", "K", "M" represents a small sample of the going literature. In a larger sample, many more letters would have been displayed.

[12] A couple of technical points. First, if we insist that K(H) be consistent, we will exclude the Newton-Leibniz calculus from Newton's celestial mechanics, thus wrecking it. The second point is a closely related one. It is sometimes overlooked how greatly a revision of K by the inclusion of H can change the original K. Think here of the impact on the science of light and all its supporting theories caused by throwing the quantum hypothesis into the mix.

Peirce's schema and AKM's too, they are hypotheses. They would be *disqualified* as hypotheses were their abducer to advance them as true. So for this kind of explanation—a widely held one in the philosophy of science—abduction cannot be inference to the best explanation (Woods 2011).

Perhaps the larger problem posed by IBE is the unruly ambiguity of "explain" compounded by variations in their theoretical developments. Riding alongside, is the uncertainty that dogs the "therefore" operator.[13] The second post-Peircean schema, the Gabbay-Woods schema, attempts to bring some clarity to this murk. Here is how it goes, informally expressed. Anyone (Bill, say) for whom E is a Peircean surprise has an ignorance problem. Bill wants to know what he doesn't know. But for the present he will settle for this. He would like to know what, if true, would remove the sting of surprise and would render E's occurrence a matter of course. He would then have some reasoned basis for making something of this hypothetical answer. He could set out to determine whether it is true. He could send it to the lab for experimental test, or he could put it to provisional use as a premiss in reasoning about things of the same subject-matter in which the ignorance problem arose in the first place.

To put this in schematic terms, let E be the surprising event and ?E the question whose answer Bill wants to know but doesn't. Let K be his present knowledge-base, K* be his present knowledge-base supplemented by new information within Bill's timely and easy grasp, let H be an hypothesis and let $\rightsquigarrow$ be the subjunctive conditional relation. C(H) is the Peircean conclusion that there is reason to suspect that H is true, and H^C is the decision to put H to whatever kind of test as seems appropriate to its subject matter. R denotes a response relation to the question posed by E. These are put to schematic work as follows:

1. ?E   [the question posed by E]
2. ~R (K, ?E)   [fact]
3. ~R(K*, ?E)   [fact]
4. H $\notin$ K   [fact]
5. H $\notin$ K*   [fact]
6. ~R(H, ?E)   [fact]
7. ~R(K(H), ?E) [fact]
8. H $\rightsquigarrow$ R(K(H))   [fact]

---

[13] Here is the Oxford-Cambridge online provision for "to explain", slightly paraphrased: to make something plain, clear or intelligible; to give a reason so as to justify or excuse something; to be the cause of something; to be the motivating factor of an action or decision; to minimize the significance of something embarrassing; or to defend oneself against verbal or oral attack.

9. H meets further conditions $S_1, \ldots, S_n$ [fact]
10. Hence, C(H)   [subconclusion, 1-9]
11. So, H^C [conclusion, 1-10]

It is worth repeating that the G-W schema applies only to the second stage of Peircean abduction. As it stands, it has nothing to say about stage one, the stage at which H was selected from among rival hypotheses to do the heavy lifting of stage two. We should also note that the G-W schema, unlike the AKM, stands mute on the conditions that fall upon H and K(H). The reason for it is partly its rejection of the AKM conditions, but mainly because those conditions will have been exposed in a principled way in the stage-one or selection-stage (tentatively and incompletely developed in Woods 2017).

Even so, our question here is whether the G-W schema offers IBE a welcome home. The answer is that it might and mightn't for particular readings of K(H), but not for all. If we agreed that when the K(H) of a successful G-W abduction is an explanatorily coherent statement of a future epistemic state, we could say that explanatory coherence is a welcome bonus. Of course, it depends on when the bonus would take effect. From this we might arrive at an explanatory hypothesis of our own.

> *Explanatory coherence as a collateral benefit*: There are cases in which an abduction is entirely successful, and such explanatory coherence as might await it is achieved in the post-H^C stage of the G-W schema. That is, it is achieved by further testing of H. There are also less dramatic cases in which the attainment of explanatory coherence is concurrent but not intrinsic to abductive success.[14]

Another less actively discussed schema for abduction is the J-J schema of Josephson and Josephson 1994.

1. D is a collection of data.
2. H explains D.
3. No other hypotheses can explain D as well as H does.
4. Therefore H is probably true.

Each premise of the schema must be *plausible* if the conclusion is to be made plausible. In *Character Evidence: An Abductive Theory*, Douglas Walton (2006) adopts the J-J schema and provides it with dialectical adaptation for legal contexts.

---

[14] Lorenzo Magnani's 2009 *eco-cognitive model* is an interesting and ambitious attempt to enlarge substantially the reach of the two schemata in view here, and in so doing remove the necessity to see them as rivals.

    i.   F is a finding or a given set of facts.

    ii.   E is a satisfactory explanation of F.

    iii.   No alternative explanation E given so far is as satisfactory as E is.

    iv.   Therefore, E is plausible, as a hypothesis. (p. 167)

We find in these approaches an emphasis on plausibility that is not to be found in the Peirce, AKM or G-W schemata. Especially striking is the difference between Walton's terminal line and the three prior termini. There is all the difference here between having reason to suspect the truth of even something you don't understand and concluding that what you do understand is plausible. It might well be that plausibility has a load-bearing role to play in sorting out different intelligible hypotheses for possible selection. Whether a stage one matter and a stage two one as well, whether a matter for abduction alone or loading-bearing in other contexts, plausibility awaits a satisfactory theoretical working up, notwithstanding its importance and the foundational work of Nicholas Rescher (1976).[15]

To conclude …

In this note, I have attempted to make some headway with the following matters:

- What are today's predominant approaches to abductive inference, and is any of them consistent with any other? I have responded with the AKM, G-W and J-J models, and have explained that each conflicts with the others. Conclusion: The logic of abduction is not yet "settled science."

- What is the relation between these models of abduction to inference to the best explanation? I have suggested that on some interpretations of "explain" IBE is consistent with the G-W model but not inherently part of it. Since the ↬ relation of the AKM model varies contextually between deductive implication and inductive strength, its tie to IBE remains obscure. A closer connection can be found between Walton's plausibility-reading of the J-J model, in which the tie to IBE is tighter. However, the plausibility-reading is at variance with some of the most important cases in abductive science. If, as I think it does, Walton's plausibility reading of the J-J model ensues from Rescher's plausibility calculus, it is incompatible with the G-W model. Since the G-W model captures utterly implausible but scientifically break-out examples, that is a reason to abandon the plausibility factor.

---

[15] For reservations see Gabbay and Woods (2005, pp. 222-238) and Woods (2014, pp. 279-287).

- The AKM, G-W and J-J models certainly don't exhaust the going approaches to abduction, chiefly those in computational logics and approaches to logic programming. I have omitted them partly for want of space, but also because of the likelihood of their soon influencing expected iterations of CT theory.
- Why does this note bother with rival approaches to epistemology? Isn't critical thinking a matter for logic and argumentation theory? My answer is that since belief-revision is an essential component of inference and decision, it is subject inescapably to epistemological constraints. To the extent that this is so, a misstep in epistemology could risk wrecking a theory of critical thinking.
- My chief purpose in this chapter has been to establish a back-channel from current theories of abductive inference and inference to the best explanation to mainstream approaches to CT theory. I hope that I have made some headway with this but, needless to say, that is something for the reader to judge.[16]

Selected readings on Peirce and abduction

The present-day literature reveals different streams of thought in bringing the Peircean paradigm into a more unified and systematic, theory. There isn't space in this chapter to adjudicate these differences. Interested readers could consult to advantage the somewhat selective works noted below, arranged in chronological order.

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

Josephson, J.R. and Josephson, S.G. (Eds). (1994). *Abductive Inference: Computation, Philosophy, Technology.* Cambridge: Cambridge University Press.

Atocha Aliseda, Atocha (1997). *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence,* PhD thesis. Amsterdam: Institute for Logic, Language and Computation.

Kapitan, Tomis (1997). Peirce and the structure of abductive inference. In Nathan Housser, Don Roberts and James van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce*, pp. 477-496. Bloomington: Indiana University Press.

Flack, P. and Kakas, A. (Eds.). (2000). *Abductive and Inductive Reasoning,* Dordrecht: Kluwer.

Magnani, Lorenzo (2001). *Abduction, Reason and Science,* New York: Plenum.

Lipton, Peter (2004). *Inference to the Best Explanation,* new revised edition. London: Routledge. First published in 1991.

---

[16] Warm thanks to J. Anthony Blair for skillful editing of an earlier draft of this chapter.

Thagard, Paul (2000). *Coherence in Thought and Action,* Cambridge, MA: MIT Press.

Gabbay, Dov and Woods, John (2005). *The Reach of Abduction: Insight and Trial,* Volume 2 of their *A Practical Logic of Cognitive Systems.* Amsterdam: North-Holland.

Walton, Douglas (2005). *Abductive Reasoning.* Tuscaloosa: University of Alabama Press.

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

Pietarinen, Ahti-Veikko (2006). *Signs of Logic,* a volume in the *Synthese* Library. Dordrecht: Springer.

Thagard, Paul (2007). Abductive inference: From philosophical analysis to neural mechanisms. In A. Feeney and E. Heit (Eds.), *Inductive Reasoning: Experimental, Developmental and Computational Approaches,* pp. 226-247. Cambridge: Cambridge University Press.

Hintikka, Jaakko (2007). *Socratic Epistemology,* 2007

Schurz, Gerhard (2008). Patterns of abduction, *Synthese,* 164, 201-234.

Magnani, Lorenzo (2009). *Abductive Cognition: The Epistemological and and Eco-Cognitive Dimensions of Hypothetical Reasoning,* a Cognitive Systems Monograph, Berlin: Springer.

Park, Woosuk (2017). *Abduction in Context: The Conjectural Dynamics of Scientific Reasoning,* a volume of Studies in Applied Philosophy, Epistemology and Rational Ethics, Berlin: Springer.

References[17]

Anselm (2009). *Proslogion.* In Ian Logan (Ed.), *Reading Anselm's Proslogion: The History of Anselm's Arguments and its Significance Today,* Burlington VT: Ashgate Publishing.

Aristotle (1985). *Categories.* In Jonathan Barnes (Ed.), *The Complete Works of Aristotle.* Princeton: Princeton University Press.

Augustine of Hippo (1888). *Tractates on the Gospel of John* 29.6. John Gibb, translator, *From Nicene and Post-Nicene Fathers,* First Series Volume 7, Philip Schaff (Ed.). Buffalo, NY: Christian Publishing Co.

Bacon, Francis (1905). *Novum Organum.* In R.L. Ellis and J. Spedding (Eds.), *The Philosophical Works of Francis Bacon,* pp. 202-387. London: Routledge, p. 292. *Novum Organum* first published in 1620.

Barringer, Howard, Gabbay, Dov and Woods, John (2012). In *Argument and Computation* 3, 143-202.

---

[17] Works prefixed with the asterisk "*" denote those that may have made some footfall in mainstream CT theories.

Flandern, Tom van (1999). Status of the NEAR challenge. *Meta Research Bulletin*, 8.

Gabbay, Dov and Woods, John. (2005). *The Reach of Abduction: Insight and Trial,* Volume 2 of their *A Practical Logic of Cognitive Systems.* Amsterdam: North-Holland.

*Harman, Gilbert (1965). The inference to the best explanation. *Philosophical Review,* 74, 88-95.

*Hintikka, Jaakko (2007). Abduction—Inference, conjecture, or answer. In Jaakko Hintikka, *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning,* pp. 38-60. New York: Cambridge University Press. First published as What is abduction? The fundamental problem of contemporary epistemology. *Transactions of the Charles Peirce Society,* 34 (1998), 503-533.

Josephson, J.R. and Josephson, S.G. (Eds.). (1994). *Abductive Inference: Computation, Philosophy, Technology.* Cambridge: Cambridge University Press.

Kakas, A., Kowalski, R.A. and Toni, F. (1995). Abduction and abductive logic programming. *Journal of Logic and Computation,* 2, 719-770.

*Kapitan, Tomis (1997). Peirce and the structure of abductive inference. In Nathan Housser, Don Roberts and James van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce,* pp. 477-496. Bloomington: Indiana University Press.

Kirby, Carmichael (1997). Letter to Olivetti. Laguna Salada Union School District, May.

Kowalski, Robert (1979). *Logic for Problem Solving: Empirical Progress of Even Truth Approximation Leading to a Challenge for Computational. New York: Elsevier.*

Kuipers, Theo A. F. (1999). Abduction aiming at modelling, *Foundations of Science,* 4, 307-323.

*Lipton, Peter (2004). *Inference to the Best Explanation,* new revised edition. London: Routledge. First published in 1991.

Magnani, Lorenzo (2001). *Abduction, Reason and Science.* New York: Plenum.

Magnani, Lorenzo (2009). *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning,* a Cognitive Systems Monograph. Berlin: Springer.

Meheus, Joke, Verhoeven, Liza, Van Dyck, Maarten and Dagmar Profijn (2002). Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In Lorenzo Magnani, Nancy J. Neressian and Claudio Pizzi (Eds.), *Logical and Computational Aspects of Model-Based Reasoning,* pp. 39-71. Dordrecht: Kluwer.

Owen, Adrian (2017). *Into the Gray Zone: A Neuroscientist Explores the Border Between Life and Death.* New York: Scribner.

Park, Woosuk (2017). *Abduction in Context: The Conjectural Dynamics of Scientific Reasoning,* a volume of Studies in Applied Philosophy, Epistemology and Rational Ethics. Berlin: Springer.

*Peirce, C.S. (1931-1958). *Collected Works.* Cambridge, MA: Harvard University Press. A series of volumes first appearing in 1931.

Peirce, C.S. (1992). *Reasoning and the Logic of Things: The Cambridge Conference Lectures of 1898,* K. L. Kettner (Ed.). Cambridge, MA: Harvard University Press.

Pietarinen, Ahti-Veikko. (2006). *Signs of Logic,* a volume in the *Synthese* Library. Dordrecht: Springer.

*Pinto, Robrt C. (2001). *Argument, Inference and Dialectic: Collected Papers.* Dordrecht: Kluwer.

Quine, W.V. (1966). *Selected Logical Papers,* New York: Random House; enlarged edition. Cambridge, MA: Harvard University Press, 1995.

*Rescher, Nicholas (1976). *Plausible Reasoning: An Introduction to the Theory and Practice of Plausible Inference.* Assen: Van Gorcum.

Rescher, Nicholas (1996). *Priceless Knowledge? Natural Science in Economic Perspective.* Lanham, MD: Rowman and Littlefield .

Russell, Bertrand (1937). *The Principles of Mathematics,* second edition. London: Allen and Unwin. First published in 1903.

Soames, Scott (2014). *The Analytic Tradition in Philosophy,* volume I *The Founding Giants.* Princeton: Princeton University Press.

Walters, Kerry S. (1994). *Re-Thinking Reason.* Albany: SUNY Press.

*Walton, Douglas (2005). *Abductive Reasoning.* Tuscaloosa: Alabama University Press.

*Walton, Douglas (2008). *Witness Testimony Evidence.* New York: Cambridge University Press.

*Walton, Douglas (2013). *Methods of Argumentation.* New York: Cambridge University Press.

*Woods, John (2004). *The Death of Argument: Fallacies in Agent-Based Reasoning,* Volume 32 of the Applied Logic series. Dordrecht: Kluwer.

Woods, John (2011). Recent developments in abductive logic. *Studies in History and Philosophy of Science,* 42, 240-244.

Woods, John (2012). Cognitive economics and the logic of abduction. *Review of Symbolic Logic,* 5, 148-166.

*Woods, John (2013). *Errors of Reasoning: Naturalizing the Logic of Inference,* Volume 45 of Studies in Logic. London: College Publications. Reprinted with corrections in 2014.

*Woods, John (2015). *Is Legal Reasoning Irrational? An Introduction to the Epistemology of Law,* Volume 2 of Law and Society. London: College Publications.

Woods, John (2017). Reorienting the logic of abduction, in Lorenzo Magnani and Tommaso Bertolotti (Eds.), *Springer Handbook of Model-Based Science,* pp. 137-150. Heidelberg: Springer.

About the author:

John Woods has done foundational work in several branches of philosophy. He is co-founder of the Woods–Walton Approach to fallacies theory, an early pioneer of the formal semantics of literary discourse and of conflict-resolution methods for the non-empirical sciences. His historical investigations have prompted the reconsideration of the structural importance for dialogue logic of Aristotle's "immature" writings in *The Organon.* In collaboration with Dov Gabbay, the logic of relevance went causal, and abduction was framed as ignorance-preserving reasoning. More recently, Woods has generalized the causal aspects of relevance to a naturalized logic of inference and legal reasoning. A new book on fiction abandons formal semantics in favour of a naturalized logic of reference and truth. Woods is the recipient of national and international awards and honorary degrees. Elected Fellow of the Royal Society of Canada in 1990, he is currently Director of the Abductive Systems Group at UBC.