Georgia Institute of Technology
SCHOOL OF PUBLIC POLICY
A Unit of the Ivan Allen College

Working Paper Series

Working Paper #2006.15

Seeing problems, seeing solutions. Abduction and diagrammatic reasoning in a theory of scientific discovery

By Michael H. G. Hoffmann
School of Public Policy
Georgia Institute of Technology

28 August 2006

Abstract:
This paper sketches a theory of scientific discoveries that is mainly based on two concepts that Charles Peirce developed: abduction and diagrammatic reasoning. Both are problematic. While *abduction* describes the process of creating a new idea, it does not, on the one hand, explain how this process is possible and, on the other, is not precisely enough defined to distinguish different forms of creating new ideas. *Diagrammatic reasoning*, the process of constructing relational representations of knowledge areas, experimenting with them, and observing the results, can be interpreted, on the one hand, as a methodology to describe the possibility of discoveries, but its focus is limited to mathematics. The theory sketched here develops an extended version of diagrammatic reasoning as a general theory of scientific discoveries in which eight different forms of abduction play a central role.

Seeing problems, seeing solutions. Abduction and diagrammatic reasoning in a theory of scientific discovery

Michael H.G. Hoffmann

Georgia Institute of Technology
School of Public Policy
D.M. Smith Building
685 Cherry Street
Atlanta, GA, 30332-0345
USA

e-mail: m.hoffmann@gatech.edu
Phone: +1-404-385-6083
Fax: +1-404-385-0504

Abstract

This paper sketches a theory of scientific discoveries that is mainly based on two concepts that Charles Peirce developed: abduction and diagrammatic reasoning. Both are problematic. While "abduction" describes the process of creating a new idea, it does not, on the one hand, explain how this process is possible and, on the other, is not precisely enough defined to distinguish different forms of creating new ideas. "Diagrammatic reasoning," the process of constructing relational representations of knowledge areas, experimenting with them, and observing the results, can be interpreted, on the one hand, as a methodology to describe the possibility of discoveries, but its focus is limited to mathematics. The theory sketched here develops an extended version of diagrammatic reasoning as a general theory of scientific discoveries in which eight different forms of abduction play a central role.

Introduction

In a recent paper, Hanne Andersen describes how physicist Ida Noddack formulated in 1934 for the first time the hypothesis that an experiment described by Enrico Fermi the same year could be explained as a case of nuclear fission (Andersen 2006). Surprisingly, the scientific community at this time did not even discuss this farsighted, and for modern physics absolutely fundamental explanation. It passed four years before Hahn and Straßmann again proposed the hypothesis that in this experiment a nucleus had split into fractions. This time, the hypothesis was immediately accepted.

There are two questions that are interesting here with regard to the general problem of the possibility, and the process, of scientific discoveries:

1. Why was Noddack's suggestion ignored?
2. How was it possible, that Noddack *saw* something in the data that others did not see?

Andersen shows that the scientific community was divided into two factions in 1934. The vast majority of scientists assumed that artificially induced "disintegration processes *had* to be one nucleus transmuting into another nucleus of almost the same size by releasing a small particle. On this model, there was no way a nucleus could divide into a few large fractions" (2006, p. 16). Noddack, on the contrary, was less constrained by physical considerations regarding a generally accepted taxonomy of possible disintegration processes. She "was an analytical chemist who had worked for years searching for the still missing elements in the periodic table" (p. 23). Therefore, she focused first of all on the chemical properties of what has been produced by Fermi's experiment. While Fermi assumed that what he produced by bombarding uranium with neutrons was the transuranic element of atomic number 93, the chemical properties the produced element displayed were different from what was expected by chemists for element 93. Andersen shows that, for Noddack, "chemical identifications clearly had much more weight in identifying the transmuted nuclei than physical expectations of possible decay series" (p. 24). This means that she felt a severe anomaly where "no anomaly was seen" by the majority of scientists; "and without a serious anomaly, there was no reason to accept a radical change of a highly successful taxonomy" (ibid.). Such a change was only possible for the physicists when by the end of 1938 also in another, well-investigated part of the original taxonomy a serious anomaly was discovered. Now there was "a gap in the taxonomy ..., ready to be filled" by the "new" phenomenon of nuclear fission (ibid.).

What I want to highlight with this example is that *seeing a solution* presupposes *seeing a problem*. Although this may sound quite trivial, it is less so if we try to formulate a theory of scientific discoveries. In this case, namely, we have to clarify what it means "to see a problem," and we have to develop a description of scientific activity in a way that the possibility of scientific discoveries can be explained within the framework of this description. This is the goal of this paper: to describe scientific activity in such a way that different possibilities of creating something "new" become visible.

Such a general description of the scientific method can be based on Peirce's concept of "diagrammatic reasoning." The central idea of this kind of reasoning is that we *see* problems when we try to represent what we know about something. The creative possibilities, on the other side, that are possible in such a situation can be specified by a distinction of different forms of what Peirce called "abduction." This concept, that Peirce introduced to describe the process of forming explanatory hypotheses, is much better known in philosophy of science than diagrammatic reasoning (see e.g., Hanson 1972 <1958>; Simon 1979; Nickles 1980b; Nickles 1980a; Grmek et al. 1981; Jason 1988; Kleiner 1993; Haaparanta 1993; Meheus & Nickles 1999; Magnani et al. 1999; Magnani 2001; Magnani & Nersessian 2002).

Both concepts, however, are problematic. As I will show in the first part of this paper, many things Peirce himself says about abduction are more confusing than helpful when it comes to *explain how* "the process of forming an explanatory hypothesis" (Peirce CP 5.171) might be possible. My thesis is that a more convincing approach can be developed when we distinguish different forms of abduction and show at which points they can arise in the process of diagrammatic reasoning. This will be the focus of the second part of this paper. The distinctions I am suggesting here are based on some further concepts Peirce developed for the first time: "hypostatic abstraction," "theorematic deduction," and the "theoric transformation" of a problem.

While Peirce introduced abduction besides deduction and induction as one of "three elementary kinds of reasoning" (CP 8.209) that are relevant for any sort of scientific inquiry (cf. CP 7.202–207; CP 6.469–476), his discussion of the other concepts mentioned above is limited to his philosophy of mathematics, or better: of creativity in mathematics. There are hardly any passages in Peirce's writings where he combines both these discussions (cf. Hoffmann 2005a, pp. 203ff.). Such a combination, however, should be interesting for an understanding of creativity in science. Therefore, I will use the terms Peirce introduced in his philosophy of mathematics for a general theory of scientific discovery.

With regard to that, the second part of this paper starts with the concept of "diagrammatic reasoning" and shows then which role the other concepts can play within this activity. Peirce defines "diagrammatic reasoning" as a three-step process: (a) constructing a representation, (b) experimenting with it, and (c) observing the results (Peirce NEM IV 47f.; cf. Stjernfelt 2000; Hoffmann 2003; 2004). The basic idea according to my interpretation is, that by representing a problem in a diagram we can not only "play" with this problem, but we can also reflect on our own cognitive and representational means by which we approach this problem. (Cf. for an example in conflict management Hoffmann 2005c.) We have to *represent* what we know—or think to know—in order to *see*, first, its limitations and, second, new possibilities. This latter, creative step will be the place where abduction and the concepts "hypostatic abstraction," "theorematic deduction," and "theoric transformation" enter the stage. Before I start, I would like to define the other concepts mentioned here because they are basic for the whole discussion.

"Hypostatic abstraction" can be defined as creating a new sign for a new object by transforming a concrete predicate into an abstract noun. Peirce gives a nice example from Moliere's *Malade Imaginaire* where a candidate for a medical degree answers the examination question "why opium puts people to sleep, by saying that it is because it has a dormative virtue" (CP 4.234, 4.463). Although this answers seems to absolutely ridiculous since instead

> of an explanation he simply transforms the premise by the introduction of an abstraction, an abstract noun in place of a concrete predicate. It is a poignant satire, because everybody is supposed to know well enough that this transformation from a concrete predicate to an abstract noun in an oblique case, is a mere transformation of language that leaves the thought absolutely untouched. I knew this as well as everybody else until I had arrived at that point in my analysis of the reasoning of mathematics where I found that this despised juggle of abstraction is an essential part of almost every really helpful step in mathematics. (Peirce NEM IV 160)

For Peirce, "hypostatization" is creating an entity based on a specific sort of abstraction. While in "prescissive abstraction" we abstract from features like color and width to define a geometrical line, in "hypostatic abstraction" we turn what can be a predicate of many things—honey is sweet, strawberries are sweet, sugar is sweet—into "a subject of thought" (CP 5.534): "sweetness" (CP 4.235). Since "hypostatization"—from the Greek *hypostasis*—is the same as "reification" in Latin, I will use here this latter term which is better established today. Whether "hypostatization" or "reification," both concepts mean creating a new thing out of what is not a thing.

Peirce highlights especially "the importance of this operation in mathematics."

> (I)t will suffice to remember that a *collection* is an hypostatic abstraction, or *ens rationis*, that *multitude* is the hypostatic abstraction derived from a predicate of a collection, and that a *cardinal number* is an abstraction attached to a multitude. So an *ordinal number* is an abstraction attached to a *place*, which in its turn is a hypostatic abstraction from a relative character of a unit of a *series*, itself an abstraction again. (CP 5.534)

"Hypostatic abstraction," or "reification," can either be the *process* of generating new signs that signify new objects, or the *product* of this process. In general, I would assume that all concepts of our languages are outcomes of reification performed at some time in the history of our cultures. This way, reification would be one of the most fundamental concepts to describe the genesis of knowledge.

Reification is also central for the next concept I mentioned above: "theorematic deduction." Peirce introduced this concept as part of a discussion that treats deductive reasoning—surprisingly—as a creative activity (cf. Hoffmann 2005b, chap. 6). He puts "theorematic deduction" in contrast to "corollarial deduction" as two forms of "necessary deductions."

> A Corollarial Deduction is one which represents the conditions of the conclusion in a diagram and finds from the observation of this diagram, as it is, the truth of the conclusion. A Theorematic Deduction is one which, having represented the conditions of the conclusion in a diagram, performs an ingenious experiment upon the diagram, and by the observation of the diagram, so modified, ascertains the truth of the conclusion. (Peirce CP 2.267; cf. CP 7.204)

The relevance of theorematic deduction again becomes obvious in mathematics. Each time when we perform a proof "by the introduction of auxiliary individuals into the argument" (Hintikka 1983 <1980>, 113, cf. 109f.), we perform a theorematic deduction. That might be a subsidiary line in geometry, or a lemma "when it comes to proving a major theorem" (Peirce EP II 96, 1901). In our terminology, the introduced "auxiliary individual" would be a reification; and this reification could either be newly created or be transferred from another context. Peirce calls the first case an "abstractional" theorematic deduction, and in the second a "non-abstractional" one (NEM IV 49; cf. Hoffmann 2005d), but it seems to be more convincing to call the first one a "creational theorematic deduction" and the second one an "analogical theorematic deduction" since we take over a reification that has been created for another purpose. Introducing the term "fission," for example, into a deductive argument in nuclear physics would be a case of "analogical theorematic deduction" since "fission" has already been used earlier in biology to describe "the division of a cell or organism into new cells or organisms, as a mode of reproduction," and in astronomy as the "breaking up of one star into two others, as postulated in one theory of the origin of binary stars" (Oxford English Dictionary).

The last Peircean term I will discuss here as interesting to describe scientific discoveries is "theoric transformation." In spite of the word's similarity to "theorematic," it is very different. While "theorematic" seems to be coined from "theorem," "theoric" refers for Peirce to the Greek "ϑεωρία" (our "theory," original meaning: "vision"). He translates this term as "the power of looking at facts from a novel point of view" (Peirce MS 318: CSP 50 = ISP 42). "Theoric" reasoning consists "in the transformation of the problem,—or its statement,—due to viewing it from another point of view" (ibid., CSP 68 = ISP 225). Thus, a "theoric transformation," or a "theoric step" in a deductive argument, means changing the perspective. We are looking at the same data, or the same representation, but in a way that opens up completely new horizons of interpretation. Peirce hints at the well-known fact that especially developing the *idea* of a proof in mathematics often depends on a "theoric" shift (CP 4.612). At several points he uses the proof of Desargues' theorem in projective geometry as an example of how a theoric transformation works (cf. my analysis in Hoffmann 2005a, pp. 170-186; Hoffmann 2005d). But we can use this concept also beyond the limits of mathematics. For example, when Aldo Leopold saw for the first time that ecological relations are not simply causal relations—remove the wolves to enlarge the deer population—but that he has "to think like a mountain" in order to being able to manage an ecosystem as a multi-dimensional configuration (Norton 2005, p. 213ff.), he performed a "theoric transformation," a "perceptual shift" (219).

Abduction

The problems we are facing with Peirce's concept of abduction can be illustrated when we simply take a look at one his best known definitions of this term (I am referring here only to what he developed after 1900 since many of his earlier considerations are insufficient, as he admits for himself: Peirce L 409, ISP 73; CP 2.102, 8.227; cf. Hoffmann 2005b, pp. 187ff.):

> Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea; for induction does nothing but determine a value, and deduction merely evolves the necessary consequences of a pure hypothesis. (Peirce CP 5.171)

The second half of this quote is not part of the definition, but an explanation for it. However, it adds something to this definition because it says implicitly that there are *only* three logical operations for Peirce, a claim that he confirms in another remark where he says that "there are but three elementary kinds of reasoning": abduction, deduction, and induction (CP 8.209). This means, however, that any form of "reasoning," or "logical operation," that is neither deduction nor induction has to be *abduction*; and that might be much more than we would expect at a first glance. According to Peirce, we find abduction not only in science as the process of "examining a mass of facts and in allowing these facts to suggest a theory" (CP 8.209), but also in any *perception* "when I so much as express in a sentence anything I see" (Peirce LOS, p. 899f.; cf. CP 5.182ff., 8.64). Even when "a chicken first emerges from the shell" and "does not try fifty random ways of appeasing its hunger, but within five minutes is picking up food, choosing as it picks, and picking what it aims to pick," this is "just like abductive inference" (Peirce LOS, p. 899f.). This unspecified broadness of possibilities to apply the concept of abduction has led to many attempts in the literature to develop classifications of different *forms* of abduction (e.g., Bonfantini & Proni 1983; Eco 1983; Shank & Cunningham 1996; Magnani 2001).

A second problem concerns the claim that abduction is supposed here to be a "logical operation" (cf. Kapitan 1992). Elsewhere I showed that this claim can only be convincing if we start from one of two possibilities: either from a very broad understanding of "logic," for example "logic" defined as "the *art of devising methods of research*,—the *method of methods*" (CP 7.59; cf. 3.618, 4.227; Fann 1970, p. 23f.); or we interpret "logical operation" as referring only to the *form* Peirce uses to describe abductive reasoning, but not to the creative act of "forming an explanatory hypothesis" (cf. Hoffmann 1999). This second possibility is evident when we look at the famous "perfectly definite logical form" of abduction the late Peirce defined as follows:

(P1) "The surprising fact, C, is observed;"

(P2) "But if A were true, C would be a matter of course,"

(C) "Hence, there is reason to suspect that A is true" (CP 5.189).

Peirce himself states in the next sentence that according to this form "A cannot be abductively inferred ... until its entire content is already present" in the second premise. That means, however, that the logical form for itself leaves the question unanswered how to get the hypothesis "A." There is no way to create a new hypothesis by a *logical* inference. "The first emergence of this new element into consciousness must be regarded as a perceptive judgment" (CP 5.192). Only *after* the new hypothesis has emerged, we can connect "this perception with other elements" (ibid.) in a logical form like the one quoted above.

However, if we accept this distinction between "*inferential aspects of abduction*" and the creative, "*perceptive aspects of abduction*" (Hoffmann 1999, p. 280), then it is hard to see how abduction can be the "the only *logical* operation which introduces any new idea" as Peirce claims in our initial quote (my emphasis). This claim could be justified only if "logical operation" means simply "method," or "strategy" (cf. Paavola & Hakkarainen 2005). Otherwise we would get a contradiction to what we discussed just a moment ago.

The third problem of this famous quote concerns the question what exactly abduction is supposed to do. In the quote, Peirce hints, on the one hand, at the "the process of forming an explanatory hypothesis" and, on the other, at introducing a "new idea." This, however, allows different interpretations. First, it is possible that the two concepts refer to two different operations, because we can form an "explanatory hypothesis" without generating a "new idea." If any *perception* is a case of abduction, we would generate a lot of "explanatory hypotheses" regarding sensory inputs though hardly any "new idea." For example, when reading a word, the word we read is a hypothesis that "explains" a sequence of letters. In this case, we form an explanatory hypothesis without introducing a new idea since the idea we associate with a certain sequence of letters exists already in our mind.

Second, it is not clear whether the "idea" we introduce is only new for us as individuals or new for our civilization. And third, an "idea" can either be what we discussed above as the result of a "reification," that is something that can be represented by a singular concept, or by a symbol, or it could be a new *perspective* on the same data as produced by a "theoric transformation." In sum, Peirce's innocent and well-known definition can describe *six very different forms of abduction* (cf. Table 1).

	If "idea" means something that can be represented by a singular symbol (based on reification)	If "idea" means the perspective on data, or on a representation (theoric transformation)
If an explanation is possible by referring to an idea that exists already in our mind	identifying abduction	Gestalt shift
If we create an idea that is new for us, but that exists already as a part of our culture's knowledge	analogical reifying abduction	analogical theoric abduction
If we create an idea that is entirely new	creational reifying abduction	creational theoric abduction

Table 1: Six forms of abduction that are possible based on CP 5.171 alone

At this point we can already see that by means of the concepts Peirce developed in his philosophy of mathematics we can specify some important distinctions that might be overlooked without this terminology. The four creative forms of abduction listed in the second and third row of the above table will be very useful later on.

Peirce himself does not specify different forms of abduction precisely enough. However, what is important is his description of the general *situation* in which abduction takes place. This situation can be determined by two points: (a) abduction starts from the *particular* as it is given in perception, and (b) it leads to something *general*. This general might be, first, a hypothesis, or a theory that is "needed to explain ... surprising facts" (CP 7.218) or, second, "a statement of fact" by which we "make intelligible" any image provided by perception (LOS 899) or, third, an activity that is driven by a general habit, or an "instinct," like the behavior of the chicken mentioned above that from early on picks the right things (ibid.).

Two things are important to note: on the one hand, that this general itself is, as Peirce once said, "in no wise contained in the data from which it sets out"; it is "entirely foreign to the data" (LOS, 898f.); and, on the other hand, that "there is no force" in this process leading from the particular to the general (CP 8.209), meaning that—maybe except in cases of perception and instinct-driven behavior—it is not predetermined what the general exactly will be by which we react to a certain particular event or situation.

Having no force on the one hand but being more successful than could be expected by chance on the other (CP 5.172, 5.591), leads to the question how our ability to infer abductively can be explained. Peirce himself hints at many places at an "instinctive" power of "guessing rightly" (cf. Rescher 1995; Fann 1970, pp. 35–38; Paavola & Hakkarainen 2005), and also at "the uncontrolled part of the mind" (CP 5.194; cf. Semetsky 2005). But since none of these approaches is sufficiently elaborated, the possibility of abduction remains at the end unexplained. In this situation, we should focus on *methods* that might, at least, *facilitate* abduction. Just this is the role "diagrammatic reasoning" can play in a general theory of scientific discoveries.

Diagrammatic reasoning

The essential feature of diagrammatic reasoning that makes this method so interesting for the description of scientific discoveries is not the three-step process I mentioned above as its definition: constructing a diagram, experimenting with this diagram, and observing the results (Peirce NEM IV 47f.). The essential point is rather that the whole process has to be performed by means—and within the limits—of a given "system of representation" (CP 4.418). When Peirce defined a "diagram" as "a representamen which is predominantly an icon of relations" that "should be carried out upon a perfectly consistent system of representation" (CP 4.418), he worked on a series of definitions to formulate the foundations of his so-called "Existential Graphs," a graphical notation of logic intended to replace algebraic notations (cf. Roberts 1973; Ketner 1996 <1990>; Shin 2002). This system of representation is characterized by a set of conventions to represent propositions and logical relations between those propositions, and a set of rules for the transformation of graphs. From a logical point of view, Peirce's Existential Graphs—at least the Alpha and Beta Part—are as sound and complete as symbolic systems of logic.

There is no question that one needs "a perfectly consistent system of representation" for representing logical implications and validity, and the same is true for representation systems in mathematics. (Although we know now that there is a principle limit of consistency *and* completeness imposed on us by Gödel's incompleteness theorems.) If we want to prove in Euclidean geometry, for example, that the triangle's inner angles sum up to 180°, we can use a parallel to the triangle's base as auxiliary line to perform the proof. The possibility of such a parallel, however, is provided only in Euclidean geometry, not in non-Euclidean ones. That is, the specific means available in this system of representation *determine* the outcome of any transformation of a Euclidean figure we perform in accordance with the rules of this system. At the same time, however, those operations are also *constrained* by the means available. If we change the system of representation—as has been done in the development of non-Euclidean geometries—we open up entirely new possibilities (cf. Hoffmann 2004). In short, a proof in mathematics is only as perfect and consistent as the representation system in which it is performed.

The consistency of representation systems is also essential when we use the concept of diagrammatic reasoning—beyond the limits of logic and mathematics—for a general theory of scientific discoveries. Not only axiomatic systems in mathematics have to be consistent, but also, for instance, the description of styles in art, the grammar we construct to understand the syntax of our everyday languages, and theories in science.

Besides consistency, there are two further features that are common to all those systems of representation. On the one hand, we need them to design a particular representation. We need an axiomatic system to construct a proof in mathematics; we need a scientific theory to formulate a hypothesis or observational statement; we need the grammar of our everyday language to formulate a normal sentence; and we need a style to draw a painting in art—although in both the latter cases knowledge about the system of representation might be mostly implicit. On the other hand, all these systems of representation play a *normative* role: in logic and mathematics we can check the validity of an inference, or a proof, by means of the rules and conventions defined by a certain system of representation; in science the plausibility of a particular statement depends on its theoretical background; in our everyday languages the correctness of expressions is determined by grammar; and in art the definition of styles is a means of classification, for example.

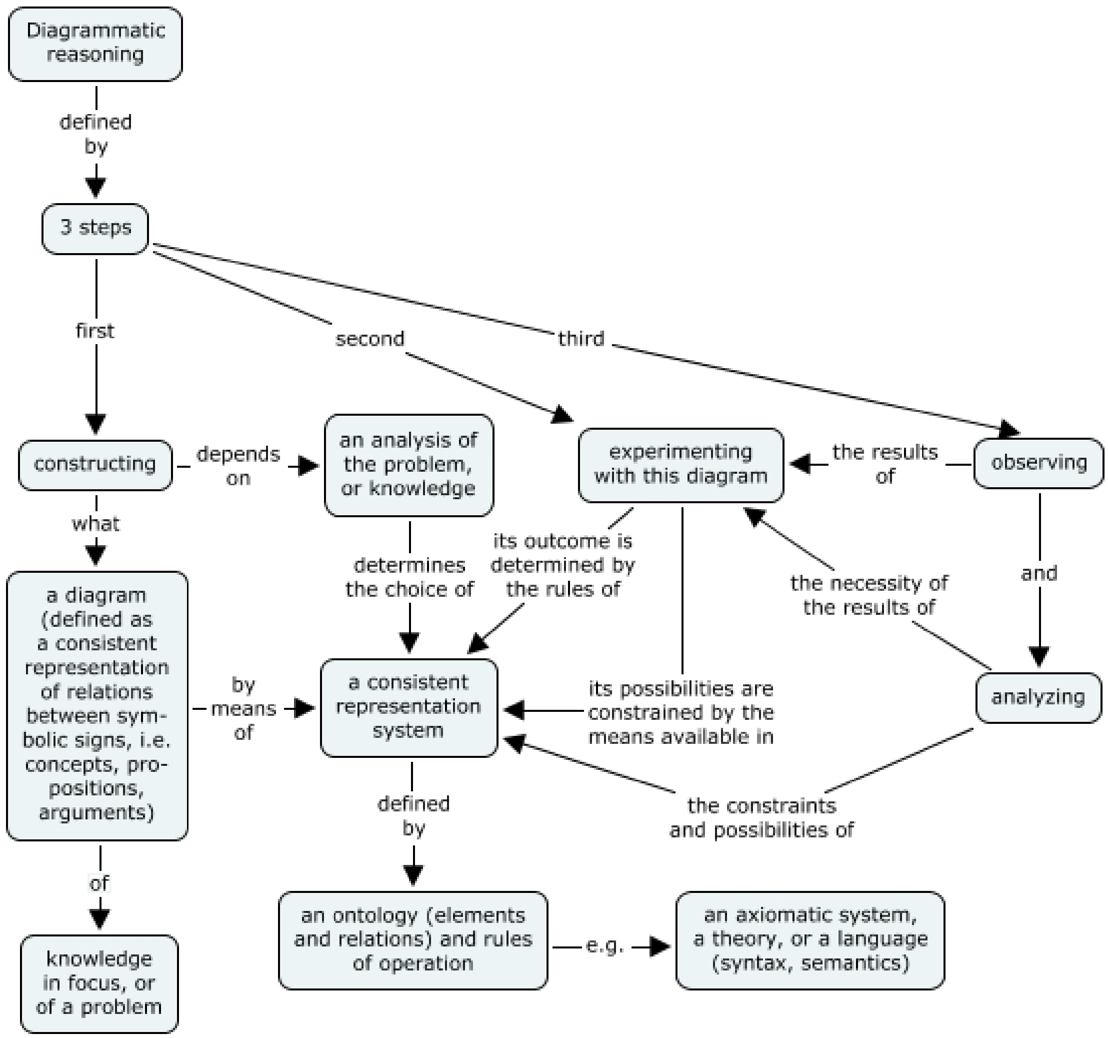

Figure 1: A definition of diagrammatic reasoning. All maps are created with IHMC Cmap tools: http://cmap.ihmc.us/

The central role of the chosen representation system for diagrammatic reasoning becomes visible when we consider its function for each of the three steps by which diagrammatic reasoning can be defined (cf. Figure 1). As this "map definition" of diagrammatic reasoning shows, the outcome of each of the three steps depends—in different ways—on the system of representation we choose in a certain situation to represent a problem, or the knowledge area we are focusing on.

The consistency and normativity of the chosen system of representation are decisive when it comes to the possibility of discoveries. Peirce once highlighted as a core idea of this *pragmatism* that all reasonings—and especially mathematical reasonings—

> turn upon the idea that if one exerts certain kinds of volition, one will undergo in return certain compulsory perceptions. Now this sort of consideration, namely, that certain lines of conduct will entail certain kinds of inevitable experiences is what is called a "practical consideration." (CP 5.9)

Such an "inevitableness" depends obviously on the normative character of the representational system in which such reasoning is performed, as is evident from mathematics: That *2* plus *2* equals *4* results from the rules and conventions of arithmetic as the chosen system of representation.

This "inevitable experience" resulting from rule-driven activity is the most important precondition of discovering something *new* by diagrammatic reasoning. As I showed elsewhere with regard to mathematics, we can distinguish two possibilities: on the one hand, the process of unfolding new *implications* of a diagram within a *given* system of representation and, on the other, the process of *developing* representational systems *themselves* that can open up new horizons and possibilities—as it has been the case in the development of non-Euclidean geometries (Hoffmann 2004).

The first case refers to the fact that we never can have a complete overview of all the implications of what we know already. Only experimentation with representations in concrete situations reveals what might already be given implicitly in our own systems of knowledge. Peirce described this case of discovering something new by saying that a diagram constructed by a mathematician "puts before him an icon by the observation of which he detects relations between the parts of the diagram other than those which were used in its construction" (NEM III 749). By experimenting upon the diagram and by observing the results thereof, it is possible, as he says, "to discover unnoticed and hidden relations among the parts" (CP 3.363).

The second case is closely connected with the problem of consistency. It might be that a rule-driven experimentation with diagrams brings to light inconsistencies or undecidable situations within our chosen system of representation. If we don't have any reason to doubt the correctness of the diagrammatic transformations we have performed, we are forced to improve the system of representation we have used.

In this situation, the *compelling* character of diagrams and the "inevitable experience" we make in diagrammatic reasoning are decisive. The results of experiments have to "stand up against our consciousness," as Hull (1994, p. 282) puts it, because only in this case can a diagram "compel us to think quite differently" (CP 1.324). It is only if we already have certain expectations concerning what should happen in processes of diagrammatization that there is a need to develop something new when those expectations are frustrated.

For a *general* theory of scientific discovery which goes beyond the limits of logic and mathematics, we can start from a critique Peirce formulated once with regard to Kant. Kant saw already that in mathematics the drawing of "necessary consequences" is possible because "the mathematician uses what, in geometry, is called a 'construction,' or in general a diagram, or visual array of characters or lines" (CP 3.560). But while for Kant the use of constructions is the criterion to distinguish been mathematics and philosophy (Kant CPR B 741), Peirce insists that there is no difference between both when it comes to the central role of "constructions," or "diagrams":

> All necessary reasoning whatsoever proceeds by constructions; and the only difference between mathematical and philosophical necessary deductions is that the latter are so excessively simple that the construction attracts no attention and is overlooked. (CP 3.560; cf.. 5.147f.)

> All necessary reasoning without exception is diagrammatic. That is, we construct an icon of our hypothetical state of things and proceed to observe it. This observation leads us to suspect that something is true, which we may or may not be able to formulate with precision, and we proceed to inquire whether it is true or not. (CP 5.162)

The connection between "necessary reasoning" and "diagrammatic reasoning" results from the normative character of the representation system used, as we saw above. However, if we try to generalize what is plausible with regard to mathematics, we are facing a serious problem. While we can hardly overemphasize the role of representations in any science, it is harder to determine what exactly the "systems of representation" in science are that we need to explain the possibility of scientific discoveries by means of diagrammatic reasoning. While in mathematics and logic we can mostly formulate a clear distinction between a "consistent system of representation" (e.g., an axiomatic system, or a notation) and a "diagram" constructed by means of this system, this is not so easy in the sciences. Of course, we could say that any science uses different—more or less formal—languages so that its representations depend on the "grammar" of those languages. But just this is part of the problem. There are quite a lot of those "languages," and they are more or less relevant for what happens in scientific activity. We would assume, for example, that it does not matter whether we formulate Einstein's theory of relativity in English or in Chinese. Sometimes, however, it matters, for example when we formulate a sociological theory that is based on ideas, or concepts, for which there is no equivalent in another culture. And it matters of course that we need for Einstein's theories the language of mathematics; sometimes mathematical theories must be developed before there is progress in physics.

The problem is, thus, that there are very different systems of representation that we often use at the same time when representing scientific knowledge, and that not all of them determine and constrain the possibilities of scientific activities in the same way as the Euclidean language determines proof possibilities in geometry. It seems to be impossible to propose a classification of all systems of representation we are using in different disciplines, and to determine in general which is relevant in which way. The problems we had to face for such an endeavor would be similar to the problems Thomas S. Kuhn faced when forced to clarify his concept of "paradigm" which is—at least with regard to its normative character—pretty similar to our concept of "representation system."

The only way I can see to cope with this problem is not to start with a "top down" classification of representation systems, but to proceed "bottom up." It should be sufficient to define the term "system of representation" only formally as that consistent set of conditions we have to presuppose to *understand* a representation, and to look at diagrammatic representations themselves guided by the question 'What kind of representation system do we need to make sense out of this expression?'. When we read a scientific text in which Sun, Moon, and Mars are called "planets," for example, a consistent representation system that could produce such an utterance would be the Ptolemaic astronomy, but not the Copernican system of representation with its classification of celestial bodies (Chen et al. 1998, p. 9).

From a general point of view, the term "system of representation" refers first of all to those conditions of understanding. Therefore, those systems can only be determined based on an analysis of concrete representations, they are *relative* to what they are supposed to determine. It depends on the concrete diagram what we have to identify as a relevant system of representation. Relevant is what we need to understand the diagram.

In order to show how, based on these considerations, the possibility of scientific discoveries can be explained, I suggest extending the definition of diagrammatic reasoning proposed in Figure 1 above. We have to continue the series of steps that are described in this definition. A complete description of what follows after the third step described in Figure 1 is presented in Figure 2. Remember, the first three steps that are already mentioned by Peirce are: constructing a diagram, experimenting with it, and observing—and analyzing—the outcome of those experiments. The continuation of the process depends, obviously, on the result of the third step. As elaborated in Figure 2, we can distinguish three possible results, each combined with a certain implication that will then again lead—at least in two cases—to further steps.

The basic ideas of this enlarged model of diagrammatic reasoning can be summarized as follows:

1. There is only one situation in which the process of diagrammatic reasoning comes—temporarily—to a halt: when we get the first possible result of observing and analyzing the outcome of an experiment, that is "the outcome does not contradict our expectations" (see Figure 2). In this situation we see that what we observe in the experiment is a necessary implication of the original diagram, and we can be happy that no contradiction occurs. There is no reason to continue the process. We learned something new, although nothing that could surprise us. It is like proving that the triangle's inner angles sum up to 180° when nobody thought about that before. It is interesting, but no reason to doubt the reliability of our cognitive and representational means.

2. Everything else besides this first possibility leads over a series of further steps finally back again to the first step, the construction of a diagram. This means, the whole process is with regard to the *activities* described in Figures 1 and 2 an endless loop that stops only when the situation mentioned in (1.) occurs.

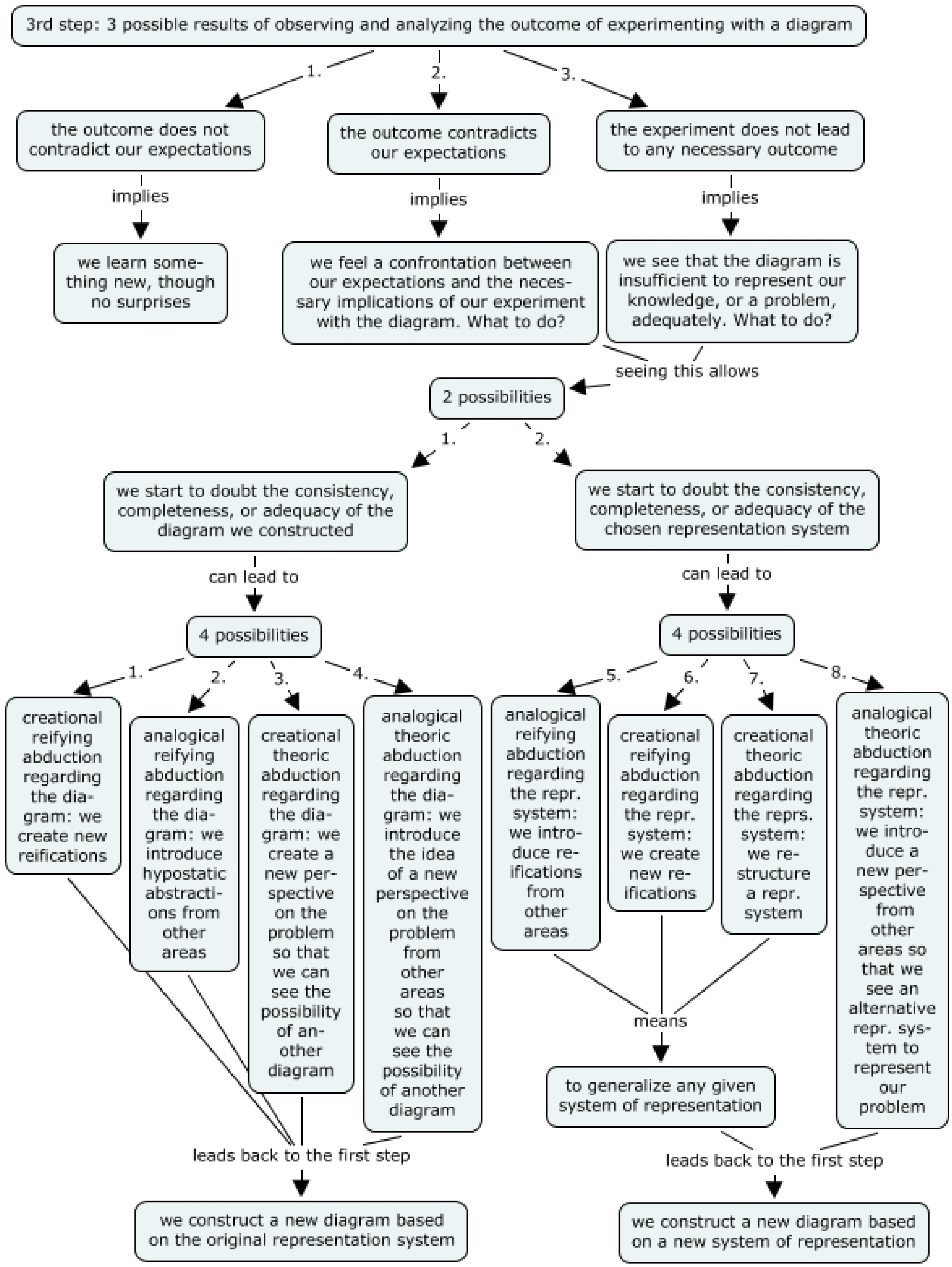

Figure 2: The process of diagrammatic reasoning, extended version, continuing the series of three steps described in Figure 1.

3. However, regarding the *representational means* available, each of these loops

continues with the first step on a more advanced level: we are either able to construct a *new diagram* by the means of the representation system that we used already from the very beginning, or we are able to construct a new diagram based on a *new system of representation*.

4. The possibility of such an advancement is the result of scientific discoveries that are differentiated in this model as eight forms of abduction (1. to 8. in the second half of Figure 2). The terminology according to which these forms are classified goes back to the concepts "reification" and "theoric transformation" as defined at the end of the introduction, and to the terms I coined based on that with regard to abduction in Table 1. Each of these forms of abduction enlarge the set of *representational*—and therefore also: *cognitive* (cf. Hoffmann & Roth forthcoming)—means available for diagrammatic reasoning.
5. The distinction of all the possibilities differentiated in Figure 2 is based on an analytical interest to separate different forms of scientific discoveries—or elements of those discoveries—as clear as possible. In practice, however, these forms will usually be intertwined, I guess.

This leads us back again to our starting point, the historical example I discussed in the beginning. It is interesting to see that the majority of physicists' "blindness" regarding the possibility of nuclear fission in 1934 can directly be ascribed to the limitations of the predominantly used "system of representation." The "diagrams and notations which were developed," says Anderson with reference to some diagrams published by Meitner in 1934, "could only represent the idea that a projectile hits a nucleus which as a result transformed into another nucleus by the emission of a particle" (Andersen 2006, p. 7). The majority of physicists was convinced by the obvious harmony between the taxonomy of disintegration processes mentioned above and "Gamow's theory of decay which precluded decay products larger than the α-particle" (ibid.). The thesis "that only particles up to the size of the α-particle could be emitted would become tacitly accepted in the whole scientific community to such an extend that the mere possibility of larger decay products [and, therefore, also the possibility of nuclear fission, M.H.] would never be mentioned" (ibid.). Background assumptions like this one form a "system of representation" in so far as they can be identified as those *conditions* that make concrete "diagrams," that is any scientific utterance formulated in this context, *understandable*.

The idea of *splitting* the nucleus into larger fragments was simply beyond the possibilities of those representational means, and since hardly anybody saw a need to doubt this well-established system of representation, there was no room for a discovery as far reaching as that of nuclear fission. Seeing a solution presupposes seeing a problem. This way, the majority of scientists was happy with what is described in Figure 2 as the first possible result of observing the outcome of experiments with diagrammatic representations. They learned something new through Fermi's experiments, but no big surprises.

For Ida Noddack, on the other side, the broadly accepted assumption of Fermi's that he produced in his experiments a transuranic element directly contradicted her expectations. These expectations, however, were based on a *different* system of representation: not the taxonomy of "disintegration processes" that fitted so well to Gamow's theory, but a system of representation that she developed over the years when working on the chemical properties of elements that were yet to be discovered. In this situation of *seeing* a severe problem regarding the consistency between her predominant system of representation and what Fermi described, she did not doubt this system, but the assumption of Fermi's. To formulate it in the language of my extended model of diagrammatic reasoning (cf. Figure 2), she started "to doubt the consistency, completeness, or adequacy" of a possible diagram that—constructed by the means of her own system of representation—should be able to include Fermi's thesis. In this situation, she performed what can be described as the third form of abduction counted in Figure 2: an "analogical theoric abduction." Since she writes in her paper that heavy nuclei might "decay" when bombarded with neutrons ("Es wäre denkbar, daß bei der Beschießung schwerer Kerne mit Neutronen diese Kerne in mehrere größere Bruchstücke zerfallen, die zwar Isotope bekannter Elemente, aber nicht Nachbarn der bestrahlten Elemente sind"; Noddack 1934), it is clear that she looks at the experiment—compared to Fermi—from "a novel point of view," that is she performs what Peirce calls a "theoric transformation." And since the idea of "decay" occurs in many areas of experience, it is not a "creational theoric abduction," but an "analogical" one. If she had already coined a new term like "Kernspaltung" (nuclear fission) for what she hypothesized based on this analogical theoric abduction, she would have performed additionally a form of "reifying" abduction.

The essential point is, however, that each of these eight possible forms of abduction leads to an enlargement of the representational and cognitive means that are then available for further processes of diagrammatic reasoning. And this again is what scientific progress is all about.

Acknowledgement

Many thanks to Professor Olga Pombo and Alexander Gerner for organizing a wonderful workshop in Lisbon in May 2006, and for inviting me to present my research. Thanks also to Nancy Nersessian, Bryan Norton, Jan Schmidt, Bob Kirkman, Jason Borenstein, and Paul Hirsch for a fruitful discussion of this paper.

References

Andersen, H. (2006). Unexpected Discoveries, Graded Structures, and the Difference between Acceptance and Neglect. In J. Meheus & T. Nickles (Eds.), *Models of Discovery and Creativity* (pp. 1-27): Springer.

Bonfantini, M. A., & Proni, G. (1983). To Guess or Not To Guess? In U. Eco & T. A. Sebeok (Eds.), *The Sign of Three. Dupin, Holmes, Peirce* (pp. 119-134). Bloomington: Indiana Univ. Press.

Chen, X., Andersen, H., & Barker, P. (1998). Kuhn's theory of scientific revolutions and cognitive psychology. *Philosophical Psychology 11*, 5-28.

Eco, U. (1983). Horn, Hooves, Insteps. Some Hypotheses on Three Types of Abduction. In U. Eco & T. A. Sebeok (Eds.), *The Sign of Three. Dupin, Holmes, Peirce* (pp. 198-220). Bloomington: Indiana Univ. Press.

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

Grmek, M. D., Cohen, R. S., & Cimino, G. (Eds.). (1981). *On Scientific Discovery: The Erice Lectures 1977*. Dordrecht Reidel.

Haaparanta, L. (1993). Peirce and the Logic of Logical Discovery. In E. C. Moore (Ed.), *Charles S. Peirce and the Philosophy of Science. Papers from the Harvard Sesquicentennial Congress* (pp. 105–118). Tuscaloosa and London: The University of Alabama Press.

Hanson, N. R. (1972 <1958>). *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*. Cambridge: Cambride University Press.

Hintikka, J. (1983 <1980>). C.S. Peirce's 'First Real Discovery' and its Contemporary Relevance. In E. Freeman (Ed.), *The Relevance of Charles Peirce* (pp. 107–118). La Salle, Ill.: Hegeler Institut.

Hoffmann, M. H. G. (1999). Problems with Peirce's Concept of Abduction. *Foundations of Science, 4*(3), 271–305.

Hoffmann, M. H. G. (2003). Peirce's "Diagrammatic Reasoning" as a Solution of the Learning Paradox. In G. Debrock (Ed.), *Process Pragmatism: Essays on a Quiet Philosophical Revolution* (pp. 121-143). Amsterdam: Rodopi.

Hoffmann, M. H. G. (2004). How to Get It. Diagrammatic Reasoning as a Tool of Knowledge Development and its Pragmatic Dimension. *Foundations of Science, 9*(3), 285-305.

Hoffmann, M. H. G. (2005a). *Erkenntnisentwicklung. Ein semiotisch-pragmatischer Ansatz*. Frankfurt am Main: Klostermann.

Hoffmann, M. H. G. (2005b). *Erkenntnisentwicklung. Ein semiotisch-pragmatischer Ansatz (Knowledge development. A semiotic and pragmatic approach)*. Frankfurt am Main: Klostermann.

Hoffmann, M. H. G. (2005c). Logical argument mapping: A method for overcoming cognitive problems of conflict management. *International Journal of Conflict Management, 16*(4), 305–335.

Hoffmann, M. H. G. (2005d). Signs as means for discoveries. Peirce and his concepts of "Diagrammatic Reasoning," "Theorematic Deduction," "Hypostatic Abstraction," and "Theoric Transformation". In M. H. G. Hoffmann, J. Lenhard & F. Seeger (Eds.), *Activity and Sign - Grounding Mathematics Education* (pp. 45-56). New York: Springer.

Hoffmann, M. H. G., & Roth, W.-M. (forthcoming). The complementarity of a representational and an epistemological function of signs in scientific activity. *Semiotica*.

Hull, K. (1994). Why Hanker After Logic? Mathematical Imagination, Creativity and Perception in Peirce's Systematic Philosophy. *Transactions of the Charles S. Peirce Society, 30*, 271–295.

Jason, G. (1988). *The Logic of Scientific Discovery*. New York: P. Lang.

Kant, I. (CPR). *Critique of pure reason* (P. Guyer & A. W. Wood, Trans.). Cambridge 1998: Cambridge Univ. Pr.

Kapitan, T. (1992). Peirce and the Autonomy of Abductive Reasoning. *Erkenntnis, 37*, 1–26.

Ketner, K. L. (1996 <1990>). *Elements of Logic. An Introduction to Peirce's Existential Graphs* (2. ed.). Lubbock: Arisbe Associates.

Kleiner, S. A. (1993). *The Logic of Discovery: A Theory of the Rationality of Scientific Research*. Dordrecht Kluwer.

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

Magnani, L., & Nersessian, N. J. (Eds.). (2002). *Model-Based Reasoning: Science, Technology, Values*. New York: Kluwer Academic / Plenum Publisher.

Magnani, L., Nersessian, N. J., & Thagard, P. (Eds.). (1999). *Model-based Reasoning in Scientific Discovery*. New York: Plenum Publishers.

Meheus, J., & Nickles, T. (Eds.). (1999). *Discovery and Creativity*. Special issues of the Journal: Foundations of Science 4(3) and 4(4).

Nickles, T. (1980a). Introduction. In T. Nickles (Ed.), *Scientific Discovery: Case Studies* (pp. 1–62). Dordrecht: Reidel.

Nickles, T. (Ed.). (1980b). *Scientific Discovery, Logic and Rationality*. Dordrecht: Reidel.

Noddack, I. (1934). Über das Element 93. *Angewandte Chemie, 47*, 653–655.

Norton, B. G. (2005). *Sustainability. A philosophy of adaptive ecosystem management*. Chicago University of Chicago Press.

Paavola, S., & Hakkarainen, K. (2005). Three Abductive Solutions to the Meno Paradox – with Instinct, Inference, and Distributed Cognition. *Studies in Philosophy and Education, 24*(3–4), 235–253.

Peirce. (CP). *Collected Papers of Charles Sanders Peirce*. Cambridge, Mass.: Harvard UP.

Peirce. (EP). *The Essential Peirce. Selected Philosophical Writings. Vol. 1 (1867-1893), Vol. 2 (1893-1913)*. Bloomington and Indianapolis 1992 +1998: Indiana University Press.

Peirce. (L). *Letter*. Available in the Peirce Microfilm edition. Pagination: CSP = Peirce / ISP = Institute for Studies in Pragmaticism.

Peirce. (LOS). *Historical Perspectives on Peirce's Logic of Science: A History of Science*. Berlin 1986: Mouton.

Peirce. (MS). *The Charles S. Peirce Papers. Manuscript Collection in the Houghton Library, Harvard University*. Available in the Peirce Microfilm edition. Pagination: CSP = Peirce / ISP = Institute for Studies in Pragmaticism.

Peirce. (NEM). *The New Elements of Mathematics by Charles S. Peirce* (Vol. I-IV). The Hague-Paris/Atlantic Highlands, N.J., 1976: Mouton/Humanities Press.

Rescher, N. (1995). Peirce on Abduction, Plausibility, and the Efficiency of Scientific Inquiry. In N. Rescher (Ed.), *Essays in the History of Philosophy* (pp. 309–326). Aldershot Avebury.

Roberts, D. D. (1973). *The Existential Graphs of Charles S. Peirce*. The Hague Mouton.

Semetsky, I. (2005). Learning by abduction: A geometrical interpretation. *Semiotica, 157*(1/4), 199–212.

Shank, G., & Cunningham, D. J. (1996). Modeling the Six Modes of Peircean Abduction for Educational Purposes. *Online*, http://www.cs.indiana.edu/event/maics96/Proceedings/shank.html (accessed July 30, 2006).

Shin, S.-J. (2002). *The Iconic Logic of Peirce's Graphs*. Cambridge MA: MIT Bradford Books.

Simon, H. A. (Ed.). (1979). *Models of Discovery and Other Topics in the Methods of Science*. Dordrecht: Reidel.

Stjernfelt, F. (2000). Diagrams as Centerpiece of a Peircean Epistemology. *Transactions of the Charles S. Peirce Society, 36*, 357–384.