2010

"Theoric Transformations" and a New Classification of Abductive Inferences

Michael H.G. Hoffmann, *Georgia Institute of Technology - Main Campus*

"Theoric Transformations" and a New Classification of Abductive Inferences

Michael H.G. Hoffmann

Forthcoming in: Transactions of the Charles S. Peirce Society. July 27, 2010

Abstract

Based on a definition of "abductive insight" and a critical discussion of G. Schurz's (2008) distinction of eleven "patterns of abduction" that he organizes in four groups, I suggest an even more comprehensive classification that distinguishes 15 forms in an alternative structure. These forms are organized, on the one hand, with regard to what is abductively inferred—singular facts; types; laws; theoretical models; or representation systems—and, on the other, with regard to the question whether the abductive procedure is selective or creative (including a distinction between "psychologically creative," as in school learning, or "historically creative"). Moreover, I argue that theoretical-model abduction—which seems to be the most important form of abduction—depends on two preconditions: first on the availability of an adequate system of representation, and second on finding a new "perspective" on a given problem, as Peirce described it with the notion of a "theoric transformation." To understand the significance of theoric transformations—especially in mathematics—it is necessary to analyze in some detail Peirce's main example for a theoric transformation: the proof of Desargues's theorem.

1 Introduction

Among the many problems posed by Peirce's concept of abduction is the question of how to determine the scope of this form of inference, and how to distinguish different types of abduction. This problem can be illustrated by taking a look at one of his best known definitions of the term:

> 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 [1903])

The second half of this quote is not part of the definition, but an explanation of 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 writes that "there are but three elementary kinds of reasoning": abduction, deduction, and induction (CP 8.209 [c.1905]). This would mean, however, that any form of "reasoning," or "logical operation," that is neither deduction nor induction has to be *abduction.* The problem that arises here is that the concept of abduction becomes very broad. According to Peirce, we do not only find abduction in science as the process of "examining a mass of facts and in allowing these facts to suggest a theory" (CP 8.209 [c.1905]), but also in any *perception* "when I so much as express in a sentence anything I see."[i] 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 II 900 [1901]).

Being "like" abduction, however, is not the same as being abduction. I think it is important to limit abduction to those forms of reasoning that generate "explanatory hypotheses," as Peirce says in the first sentence of the quote above. There has been a debate recently whether there are also "non-explanatory" forms of abduction. Following Gabbay & Woods (2005), Lorenzo Magnani (2009) discusses "instrumental abduction" as a form of "non-explanatory abduction" (Ch. 2). As an example, he hints at reverse mathematics where axioms are introduced that are justified only by the fact that they can be used to prove a target theorem. At first glance, these axioms do not "explain" anything, but they are nevertheless created as newly formed hypotheses that are instrumental for the theorem in question (p. 72). Another example is the assumption of implausible hypotheses for purely instrumental reasons, like Newton's

assumption of "action at-a-distance … that allows the gravitational theory to predict observations that it would not otherwise be able to predict" (p. 77). Since it seems plausible to assume that we would accept as explanans only those propositions, or sets of propositions, that we can accept as being true, coming up with implausible hypotheses would indeed be a case of non-explanatory abduction.

In response to Magnani's suggestion to enlarge the "orthodox Peircean *explanatory* view" of abduction by including "non-explanatory and instrumental abduction" (p. 9), Paul Thagard (2010) suggests that "non-explanatory abduction" should not be called "abduction," but something else. I agree. Since terminological questions are questions of convenience, not of truth, I would argue that we get into serious practical troubles if we don't limit abduction to "forming an explanatory hypothesis," as Peirce suggested. If we broaden the meaning of abduction so that it includes everything someone might come up with, then we could end up counting any daydream as an abductive inference. What is the criterion by which the decision could be justified to count "instrumental abduction" as abduction, but not any creation of "ideas"?

An important part of this discussion should be the question of how we can know when the goal of performing an abductive inference has been achieved. Elsewhere, I discussed this question as the need of a stopping rule (Hoffmann, 2010). Whatever our definition of abduction might be, abductive reasoning can only be delimited from non-abductive reasoning if we have a clear understanding of when exactly the function of abductive reasoning has been fulfilled. Thus, I argued that a sufficient understanding of abduction must include a definition of an "abductive insight" that marks the point at which the goal of abductive reasoning is achieved. The definition, however, that I proposed in this context (p. 45) needs to be revised if we assume—as I do—that the goal of abduction is an explanation. In this case, a definition of abductive insight must be formulated—to avoid circularity—so that it clarifies, at the same time, our understanding of what an "explanation" is. I propose, thus, that the search for an explanatory hypothesis can be stopped when an abductive insight has been achieved, which I define as follows:

An insight resulting from the creation of an explanatory hypothesis is the experience that what someone created in abductive reasoning fits into a system of beliefs—or provides such a system—that fulfills three conditions: (1) each one of the beliefs in the system is acceptable to the person experiencing the insight; (2) the system as a whole satisfies for this person the need of understanding either a particular phenomenon or a general regularity so that the phenomenon or the regularity can be perceived as plausibly connected to this system of beliefs; and (3), in order to show that this connection between the system of beliefs and the phenomenon or regularity is indeed plausible, it must be possible to represent this connection in the form of an acceptable argument whose conclusion is a proposition describing the phenomenon or regularity.

This definition of abductive insight combines several elements that have been discussed in the long and complex debate on "scientific explanation" in the philosophy of science. The third condition includes both Hempel-Oppenheim's "Covering-Law" or "Deductive-Nomological" model of explanation and Salmon's "Causal" or "Statistical Relevance" model as possibilities (Hempel & Oppenheim, 1948; Salmon, 1984), but also "mechanistic explanations" as described by Bechtel and Abrahamsen (2005) for biology;[ii] it even goes beyond these well-known approaches to explanation because can additionally include arguments that are based on expert opinion or analogy and whatever can be accepted as an argument (including visual arguments). The second condition, on the other hand, uses ideas from van Fraassen's pragmatic model of explanation and Friedman's unification model (van Fraassen, 1980; Friedman, 1974).

The main difference to this well-known tradition, however, is that the understanding of explanation developed in my definition is clearly subjective and contextual. The non-objectivity and contextuality of explanations shouldn't be a problem, however, since explanations are considered here only as parts of abductive insights which are, obviously, personal experiences. Because we need the concept of abductive insights to formulate a stopping rule, and since I am convinced that the goal of abduction is the

creation or selection of an explanatory hypothesis, it is sufficient to define explanation in the subjective and contextual way proposed in the definition above.

Based on this definition we can exclude the process of coming up with implausible assumptions as a form of abduction because assumptions that are implausible for the person that tries to hypothesize an explanans violate the first condition. However, if Newton in the example used by Magnani would have accepted "action at-a-distance" at least as a possibility, he himself would have accepted his reasoning as providing an explanatory hypothesis, and hence as abduction. The fact that others and later generations, and maybe Newton himself, did not accept this assumption would only mean that for those people the process of looking for an acceptable explanans did not come to an end with the assumption of "action at-a-distance." For them, the search for "an acceptable argument whose conclusion is a proposition describing the phenomenon or regularity" of gravitation would still be open.

Magnani's example of formulating axioms that are instrumental for proving certain theorems in mathematics, on the other hand, would be a case of explanatory abduction according to my definition. All three conditions seemed to be fulfilled. The same would be true for the case of perception and the behavior of the chicken mentioned above. *If* we choose to describe both in a way that *can* be represented as an argument whose components are beliefs, then both, perception and the behavior of the chicken, can be conceptualized as cases of abductive inferences.

This way, the proposed definition of abductive insights does not only deepen our understanding of abduction by providing a stopping rule for the search of explanatory hypotheses, it also provides a clear criterion to distinguish abduction—defined as "the process of forming an explanatory hypothesis," as Peirce writes—from other forms of creating something.

Anyway, the unspecified and wide scope of abduction has led to many attempts to distinguish different *forms* of abduction.[iii] The need of a comprehensive overview of possible forms of abductive inferences became even more pressing over the last

decades when Peirce's concept experienced an impressive renaissance in a variety of disciplines from computer science[iv] to theories of learning and creativity in cognitive and educational sciences,[v] and many others.[vi]

The objective of the following considerations is, first, to review and discuss critically the latest proposal for a comprehensive classification of different forms of abduction which G. Schurz (2008) provided. Based on the crucial importance of what Schurz discussed as "theoretical-model abduction," I will then argue that a sufficient understanding of this form of abduction requires a better understanding of its preconditions. As such, I think it is unavoidable to reflect, on the one hand, on the need of systems of representation that provide the means necessary to represent theoretical models and, on the other, on the significance of "shifting the perspective" that Peirce introduced for the first time as "theoric transformation." Since Peirce's considerations are scattered across a few, late manuscripts, it will be necessary to present especially his main example for a theoric transformation—the proof of Desargues's theorem, which he calls the "ten point theorem"—in a more systematic manner. The reflection on the need of representational systems will lead, in the third part, to an enlargement of Schurz's classification by what I call "meta-diagrammatic abduction." My overall goal is to provide a list of possible forms of abduction that is more comprehensive than what is available at the moment, and to argue for its completeness.

2 Schurz's "Patterns of Abduction"

A very useful classification of "Patterns of Abduction" has recently been proposed by G. Schurz, 2008). His distinction of different kinds of abduction is organized by the question of what kind of entity is abductively inferred. I will use Schurz's list as a starting point, to which I add another form of abduction that has not been mentioned so far, and a few more distinctions. The result—a classification of 15 forms of abduction for which I suggest a new organization—is summarized at the end in

Section 4 in Table 1. Schurz, by contrast, distinguished the following four main kinds of abduction:[vii]

1. In *factual abduction* we select either an observable or an unobservable singular fact as a possible explanation of a given piece of evidence—based on a known law that connects this fact with the evidence—or we select an uninstantiated type of fact in what Schurz calls "first-order existential abduction" (208). While Schurz distinguishes thus three sub-forms of factual abduction—observable, unobservable, and first-order existential—I think it is more appropriate to distinguish only two forms which are significantly different: factual abduction and what I would call "type abduction" (a term that, first, seems to me more illuminating than "first-order existential abduction" and that can be used, second, to refer to both "first -order existential abduction" and theoretical concept abduction which Schurz discussed as part of his fourth group, p. 226). The reason for this distinction is that it seems to be more significant to distinguish abductively hypothesized *types* of facts from *singular* facts than observable singular facts from unobservable ones. I think it is important whether we explain the imprints of sandals on an otherwise empty beach either by "someone was walking on the beach" (type) or by "my friend Paul walked here" (singular fact). I find it less important, from a methodological point of view, whether an "observable" person like Peter left the footprints in the sand or a no longer observable dinosaur left imprints on a beach that is now a rock.

Schurz discusses only historical facts as "unobservables," but not those things that are too small to be observable, as certain physical, chemical, or biological entities. While he takes the latter as "theoretical facts" into account, describing them as "in principle" unobservable, he argues that the abduction of theoretical facts should not be discussed as part of factual abduction but "under the separate category of 'theoretical-model abduction'." His reason for this decision is that the corresponding process of hypothesis creation is "usually not driven by simple implicational laws, but by a quantitative theory, and the abduced theoretical fact corresponds to a theoretical

model of the observed phenomenon" (209). However, this argument loses its foundation when our classification of patterns of abduction—as proposed by Schurz himself—is only determined by the question of what types of *things* are abductively inferred. If this is our criterion, then it does not matter what "drives" abduction. Moreover, if the obvious difference between "theoretical facts" and "theoretical models" becomes blurred as in Schurz's argument, then we are losing something that is crucial: the distinction between the elements of a structure and the structure itself. Theoretical facts may be elements of models, but they are not themselves models.

Peirce discussed the creation of new types under the heading of "hypostatic abstraction." "Hypostatic abstraction" can be defined as creating a new sign for a new general type by transforming a concrete predicate into an abstract noun. That is, 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 [c.1905]): "sweetness" (CP 4.235 [1902]).[viii]

Peirce described this kind of hypostatic type abduction with an 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" (as Peirce paraphrases the example in CP 4.234 [1902]; see also 4.463 [1903]). At first sight this answer seems to be ridiculous since instead

> of an explanation he [i.e., the candidate] 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 [1903])

The last point is indeed crucial. We have to keep in mind that it took the ancient mathematicians about 300 years to form a concept like "incommensurability," that is a new type or theoretical concept, although the property of incommensurable lengths has been known all along. It is by means of those hypostatic abstractions that we create new objects that we can then use as *means* for extending our knowledge; just as we needed the concept of incommensurability to create the concept of a "real number," and this again for all the other arcane number concepts we use today.

Regarding the possibility of creating new "facts" let me hint at the recent discovery that the long known disease Bovine spongiform encephalopathy (BSE), commonly known as mad-cow disease, is believed to be caused by a specific type of misfolded protein called a prion. Although common wisdom declared it impossible that proteins can carry infections, just this is now generally accepted knowledge, providing thus an example of what we might call creative—not selective—fact abduction. The singular fact is the folding prion, and what it explains is BSE in cattle and a new variant of Creutzfeldt–Jakob disease in humans.

2. In *law-abduction* "both the evidence to be explained and the abduced hypothesis is an implicational law, and the abduction is driven by one (or several known implicational laws" (211). Schurz provides as an example the abductive inference from the "background law: Whatever contains sugar tastes sweet" and the "empirical law to be explained: All pineapples taste sweet" to the conjecture "All pineapples contain sugar" (212). Both factual abduction and law abduction are "mainly selective," according to Schurz's classification, meaning that neither the laws nor the facts that are abductively inferred are newly created.[ix] In the pineapples example, this selective character of law abduction becomes visible when we assume that we have several competing background laws. In this case one has to select "the most 'plausible' one."

I would argue, however, that the vagueness of "mainly selective" should be replaced by a clear distinction: there are forms of factual and of law abduction that are "selective" and others that are "creative." While we select in the examples of the

footprints and the pineapples, it is also possible that we *create* a fact like the prion, or a type or law.

3. *Theoretical-model abduction* is the first form of abduction that is mainly creative according to Schurz, although the abductive creation of a new theoretical model that describes the causes of a phenomenon is driven "by an already *given* theory" (213) and does not introduce any new concept (216). While in the first two forms of abduction the objective is to select a fact or a law from a variety of alternatives—which corresponds to the well-known interpretation of abduction as "inference to the best explanation"[X]—the goal of theoretical model abduction is "to find just *one* plausible theoretical model which allows the derivation of the phenomenon to be explained" (213).

Schurz hints as an example for "theoretical-model abduction" at Archimedes's explanation of the phenomenon that some objects are swimming on water while others are sinking by means of the buoyancy model. "Archimedes' ingenious abductive conjecture was that the amount of water which is supplanted by the swimming or sinking body tends to lift the body upwards, with a force f_w which equals the weight of the supplanted water ... If this force is greater than the weight of the body (f_B) the body will swim, otherwise it will sink" (213). Although this abductively inferred model was something new at the time, Archimedes did not introduce any new concepts, and he created the model entirely in terms of the already given theories of mechanical and gravitational effects. This means that theoretical-model abduction is an independent form of abduction. We can explain a phenomenon or a regularity by creating a new theoretical model without using any new facts, types, or laws in this new model.

4. *Second-order existential abduction*. Whereas, according to Schurz, the three previously listed forms of abduction "are driven by *known* laws or theories, and hence, they work within a *given conceptual space*," this final group of forms that we find in Schurz's classification abduces "an at least *partly* new property or kind concept governed by an at least partly new theoretical law" (216). A first sub-group of

"second-order existential abduction" includes "micro-part abduction" in which an only partly new concept is abductively inferred. Here, "a hypothesis about the microscopic composition of observable objects" is created by extrapolation from these observable objects, and it is hypothesized that the "microparts obey the *same* laws as the observable macroscopic objects." As an example, Schurz hints at the ancient hypothesis that processes like the dissolution of sugar in water can be explained by "atoms" which—although unobservable—"obey the same mechanical laws as macroscopic bodies" (216).

A second form of "second-order existential abduction" is "analogical abduction" according to Schurz. Here not only a partially new concept like "atom" is abductively inferred, but also "new laws which connect this concept with given (empirical) concepts, in order to explain the given law-like phenomenon. The concept is only partly new because it is analogical to familiar concepts, and this is the way in which the concept was discovered. So analogical abduction is *driven* by analogy" (217).

Again, Schurz's terminological suggestions do not seem to me particularly plausible. According to our usual understanding of "analogical," analogical abduction should refer only to a certain *method* of creating hypotheses, but not to the question of *what* is created by this method. And the abduction of "microparts" is no less analogical than the abduction of the laws he hints at.

For Schurz, the "most fundamental kind of conceptually creative abduction" is what he calls "hypothetical (common) cause abduction." Without presupposing "any background knowledge except knowledge about those phenomena which are in need of explanation," this form of abduction infers "a *new unobservable entity* (property or kind) together with *new laws* connecting it with the observable properties" (218). Regarding this most creative form of abduction, Schurz proposes a distinction between "scientific common cause abduction" and abduction that hypothesizes an unobservable cause purely based on speculation. The idea seems to be that there has to be at least a minimum of "justificational value" in any abduction (204) so that

"speculative abduction" that tries to explain just one phenomenon by one abductively inferred cause has to be excluded as unscientific. To do so, Schurz proposes a "minimal adequacy criterion for second-order abductions": they have to provide "a *causal* or *explanatory unification"* by explaining "*many intercorrelated* but *analytically independent* phenomena" (219).

The main function of this last distinction seems to be the exclusion of non-scientific explanations, as in explaining a singular observation like "John got ill" by a hypothesis like: "Some power wanted that John gets ill, and whatever this power wants, happens" (219). The criterion of causal unification, however, is hardly sufficient to exclude such purely speculative abduction. We only have to look at theological ideas that do not refer to individual events like "John got ill" but to general rules like the rule that the evil gets punished while the "God-fearing" will prosper. In this example we find an extraordinary high degree of "unification" and "predictive power" in spite of a lot of speculation. This one "cause" can obviously explain "*many intercorrelated* but *analytically independent* phenomena." On the other hand, when Ida Noddack speculated in 1934—against common wisdom, and anticipating a hypothesis that gained broader support only four years later—that the "Possible Production of Elements of the Atomic Number higher than 92" which Enrico Fermi believed he had produced by bombarding uranium with neutrons might in fact be a case of nuclear fission, she obviously performed a "speculative abduction"—trying to explain just one phenomenon by one abductively inferred cause—that was nevertheless highly significant from a scientific point of view (see Andersen, 2009).

Another objection against Schurz's classification is that the term "hypothetical-cause abduction" can be applied to each form of abduction he distinguished up to this point: factual, law, and model abduction, all these forms of abduction are supposed to fulfill the function of explaining something by a "hypothetical cause." Remember, whatever kind of classification we propose, each form of abductive reasoning should fulfill the Peircean criterion of providing an "explanatory hypothesis."

The main problem of Schurz's classification of "patterns of abduction" seems to me that his list is determined by three criteria that simply do not allow for one, coherent list: first, the question of *what* is abductively inferred; second, the degree of creativity involved; and third, what kind of background "drives" the creation of explanatory hypotheses. In contrast to Schurz's approach, it should be less confusing to organize different forms of abduction in a table, not in a list, because in a table we can combine different criteria. Thus, I will propose below a distinction of five different things that can be abductively inferred—facts (observable or unobservable); types; laws; theoretical models; and systems of representation—where each of them can either be selected from an already given database or be newly created.

However, before I develop these new ideas regarding classifying forms of abduction, let me go back to what Schurz discusses as "theoretical model abduction." When it comes to the question of what can be abductively created, this form of abduction seems to me more important than what Schurz highlights as "second-order existential abduction."

3 Meta-diagrammatic abduction and theoric transformation

Schurz describes "theoretical-model abduction" by the task of

> finding theoretical (initial and boundary) conditions which describe the causes of the phenomenon in the theoretical language and which allow the mathematical derivation of the phenomenon from the theory. Formally, these theoretical conditions are expressed by factual or lawlike statements, but their semantic content corresponds to what one typically calls a theoretical model for a particular kind of phenomenon within an already given theory, whence I speak of 'theoretical-model abduction'. (213)

However, if we restrict the goal of theoretical-model abduction to those explanations in which the explanandum can be mathematically derived from the explanans, then we could only count deductive-nomological and statistical explanations as explanations. According to the definition of explanation that I provided above in connection with my

definition of abductive insight, any relation between an explanans and an explanandum can be accepted as an explanation as long as it is possible to represent this relation in form of an argument that is acceptable to the person accepting the explanation.

More important, however, than the questionable limitation to "mathematical derivation" is Schurz's observation that theoretical models represent relations between conditions, or causes, and the phenomenon to be explained. I would say the essential point in theoretic-model abduction is indeed that not individual facts, theoretical concepts, or laws are abductively inferred but the *structures* in which representations of these things are *related*.

According to my definition of explanation, this structure is always the structure of an argument—or, more precisely, a structure that *can* be represented as an argument, including visual arguments like a cartoon that provides certain "reasons" for a certain "conclusion" in form of a picture (Blair, 2004). This means, it is just such an argument that can be conceived as a theoretical model. A person can "understand"—to use the formulation of the second condition of abductive insight introduced above—either a particular phenomenon or a general regularity if he or she can perceive this phenomenon or regularity as plausibly connected with a certain system of beliefs, that is, a set of assumptions that can be represented as the premises of this argument.

This way, theoretical-model abduction would be the same as imagining the structure of an explanatory argument. Sometimes an explanation requires only hypothesizing a new model, that is, a possible argument in which well-known facts, types, and laws are merely structured in a new way. But if an explanation is only possible by hypothesizing a new fact, theoretical concept or law, then this new element must nevertheless be embedded in the structure of a possible argument to generate an abductive insight. This means, there is no abduction whose result would not be part of what can be represented as a theoretical model.

Based on this argument we can conclude that theoretical-model abduction is indeed a precondition for any abduction. Whatever we create as a hypothetical explanans, it has

to be integrated in a certain theoretical model, that is, a possible argument.

An important point is now that theoretical-model abduction itself is obviously dependent on two preconditions. The first one is that we need a system of representation that provides the means necessary to represent a model, that is, a language that provides a certain ontology, semantics, and syntax. Such a system of representation provides the means to represent facts, general types of entities or theoretical concepts, and laws, and it defines certain rules that determine how to operate with representations and how to transform them. This is important because it shows us that we need to add to our list of things that can be abductively inferred systems of representation. For example, the development of non-Euclidean geometries out of Euclid's geometry can be seen as the creation of new representation systems that provided the means for entirely new sets of theorems and proofs.

Since any operation within a certain system of representation can be described as a case of what Peirce introduced as diagrammatic reasoning (Hoffmann, 2004, in preparation), I call the process of creating new representation systems "meta-diagrammatic abduction." Since completely new theoretical models are possible when we develop new systems of representation, "meta-diagrammatic abduction" should be counted as an independent fifth form of abduction besides fact abduction, type abduction, law abduction, and theoretical-model abduction.

The second condition for theoretical-model abduction is that we have to find a certain *perspective* on the problem in question. Whatever we are approaching, we have to "frame" the subject of our attention in some way. This does not only mean that we have to choose a certain system of representation to represent the problem and its explanation, but also a certain perspective. Our vantage point determines the set of available theoretical models. It is possible to generate new models simply by shifting the perspective on a problem.

To understand the significance of finding an adequate perspective, and of shifting between possible perspectives, as a precondition for theoretical-model abduction, it is

worthwhile to have a look at another concept that Peirce introduced, so it seems, for the first time: his concept of a "theoric transformation." Peirce coined this term only around 1907, as far as I know. Its significance, however, has not yet been appreciated in the literature.[xi] For Peirce, "theoric" refers to the Greek *"theoria"* which originally meant "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 an 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 importance of developing first the *idea* of a proof in mathematics before the actual proof can be performed, and calls this a "theoric step."[xii]

In MS 318, a long manuscript titled "Pragmatism" and written in 1907, Peirce introduced the concept "'theóric' deduction" with a hint at the demonstration of Desargues's theorem (which he calls here and at other places the "ten point theorem").[xiii] With regard to Figure 1, the theorem can be formulated as follows: Given two triangles $X_1Y_1Z_1$ and $X_2Y_2Z_2$, if the rays X_1X_2, Y_1Y_2, and Z_1Z_2 intersect in O, then the intersection points A, B, and C—A being the intersection of Z_1Y_1 and Z_2Y_2 ; B the intersection of X_1Z_1 and X_2Z_2; and C of X_1Y_1 and X_2Y_2—form one straight line ABC.

As Peirce emphasizes in 1909, the surprising fact is that these three points always *"will lie on one ray,"* whatever the lengths of the lines may be (NEM III 871). In the earlier manuscript, Peirce wrote:

> Thus, nothing whatever is said in the enunciation about any definite lengths or proportion of lengths. Yet, for many years, the only known way of demonstrating the theorem was by a tedious calculation of lengths. Since this introduced an idea to which, indeed, no doubt attached, yet which was in nowise involved in the premisses, the

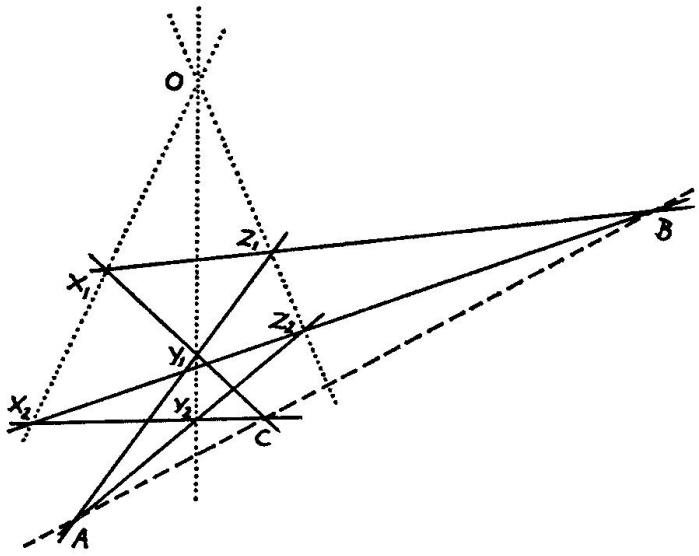

Figure 1: Peirce's diagram of Desargues's theorem (NEM II 212)

> demonstrations was of the kind called theoric,[xiv] though the new point of view was ill chosen and even, to my apprehension, illogical, since geometrical measurement of length must depend upon this very ten-point theorem, so that the latter cannot logically be made to depend upon the former. But after many years during which countless numbers of mathematicians must have asked themselves whether no better demonstration could be found, and asked in vain, von Staudt had this idea, equally simple indubitably, and penetrating, that the three rays in the plane passing through the arbitrary chosen point [that is, *O* in Figure 1; M.H.], might be regarded as a perspective representation of three rays through the same point. The moment this manifestly true suggestion is accepted, the demonstration follows from it with the aid of a few obvious definitions familiar in the theory of perspective by the easiest corollarial reasoning. Everything is corollarial except the single idea that the plane figure is a projection of a figure in three-dimensional space. That is certainly not corollarial, since there is nothing in the problem to suggest it,—no reference to a third dimension. (Peirce, MS 318 [1907]: CSP 52-53 = ISP 42-43)

Although it might be hard to "see" the possibility of such a theoric transformation, or

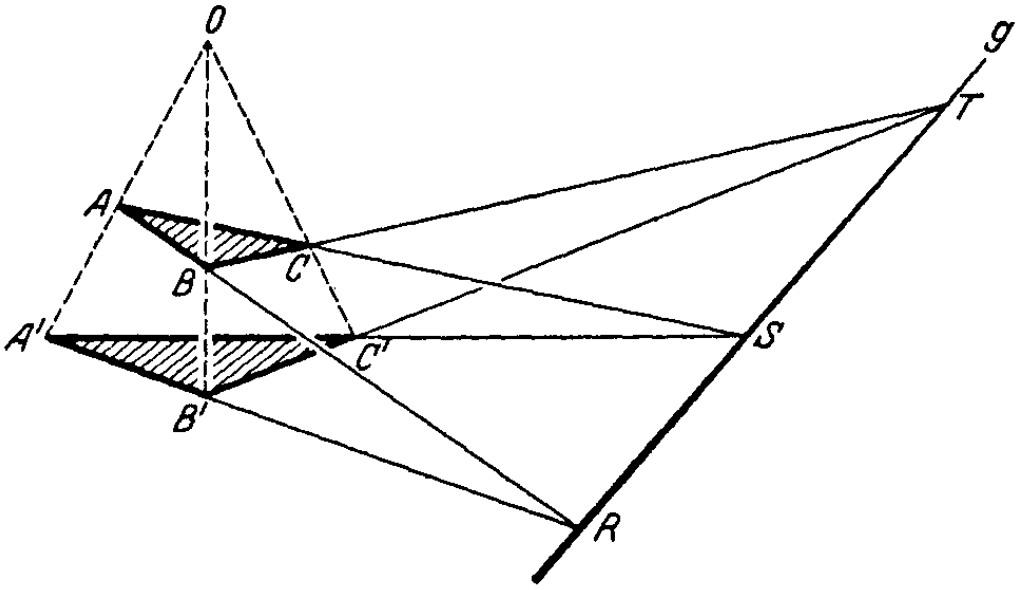

Figure 2: A representation of Desargues's theorem that is equivalent to Figure 1, but more easily to perceive as a three-dimensional representation (from Hilbert & Cohn-Vossen, 1973 <1932>, p. 107).

shift of perspective, in Figure 1, the logical necessity that A,B, and C lie on one ray is immediately evident when we look at an equivalent representation of the theorem as given in Figure 2 (taken from David Hilbert und Stefan Cohn-Vossen's book *Anschauliche Geometrie,* Hilbert & Cohn-Vossen, 1973 <1932>, 107).[XV]

If we perceive point *O* as the top of a pyramid which is cut by two planes, we see immediately that the intersection of two planes in a space *must* form a straight line. That means, the "theoric step" of perceiving a two-dimensional diagram as a 3D figure transforms Desargues's theorem—at least intuitively—into a triviality.

It should be clear that "theoric transformations" are relevant 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 "perceptual shift" (219) or, in the words of Peirce, a "theoric transformation." In the literature on policy, conflict, and communication the same

phenomenon has been discussed under the heading of "framing" and "reframing."[xvi]

What this example shows is that finding an adequate perspective on a problem can be crucial. The surprising phenomenon that the three points mentioned in Desargues theorem always lie on one ray can be explained when well-known geometrical facts and theorems can be embedded in an argument, a theoretical model, whose structure is determined by a certain perspective on the problem. The more general point that can be derived from this example is that every explanation realizes some perspective. There is no explanatory argument that is not based on a certain argumentative structure, even when we don't realize it because this structure is simply evident—as in the case when we explain the footprints on the beach by the assumption that a human being went there.

4 A new classification of abductive inferences

While I focused so far only on the things inferred by abduction—facts, types, laws, theoretical models, and representation systems—it should be important to distinguish three different *procedures* by which we can infer all these things. We saw already that Peirce conceptualizes even perception as a case of abduction. It seems to be clear, however, that we hardly create anything "new" in perceiving something, although we could say that we are "forming an explanatory hypothesis" with regard to what is in front of our eyes. Thus, if perception is accepted as a form of abduction, we need to make a distinction between *creating* an "explanatory hypothesis" and *selecting* one from a given database. For example, when reading a word, the word we read is a hypothesis that "explains" a perceived sequence of letters. In this case, we are selecting an explanatory hypothesis from the set of words we already know without creating a new one; we associate a certain sequence of letters with a hypothesis that exists already in our mind.

This way, it should be useful to distinguish between *selective abduction*, which uses an

idea as explanation that exists already in our mind, and *creative abduction* (see already Magnani, 2001). The latter, however, should again be divided into those forms of abduction that create something that is only new for us as individuals, and those that are new for our civilization. Margaret Boden, 2004 <1990>) captured this alternative by distinguishing between "psychological" and "historical creativity": There are "P-creative" ideas that are new for the person who comes up with it, especially in "learning by discovery," and "H-creative" ideas no one else had before in the history of mankind.

If we combine all the distinctions discussed so far, we can distinguish 15 very different forms of abduction which can be organized and named as suggested in Table 1.

	If an explanation is possible by a hypothesis that exists already in our mind	If we create a hypothesis that is new for us, but exists already in our culture	If we create a hypothesis that is historically new
If the "explanatory hypothesis" includes a new fact	Selective fact abduction (as in explaining a disease based on what we know)	P-creative fact abduction (as in explaining a disease by a cause that is new to us)	H-creative fact abduction (as in explaining a disease by a new cause)
If the "explanatory hypothesis" includes a new type or concept	Selective type abduction (as in reading a word, or explaining a footprint by "a human being went here")	P-creative type abduction (as in learning the concept of inertia or energy by discovery)	H-creative type abduction (as in creating the concept of inertia)
If the "explanatory hypothesis" is a law	Selective law abduction (as in explaining an event by a stereotype)	P-creative law abduction (as in learning a law by discovery)	H-creative law abduction (as in Boyle's discovery of the law named after him)
If the "explanatory hypothesis" is a theoretical model	Selective model abduction (as in framing an issue from a habitualized perspective)	P-creative model abduction (as in learning an explanatory model by discovery)	H-creative model abduction (as in Archimedes's creation of the buoyancy model)
If the "explanatory hypothesis" is based on a new system of representation	Selective meta-diagrammatic abduction (as in deciding whether a proof should be performed algebraically or geometrically)	P-creative meta-diagrammatic abduction (as in discovering that a geometrical problem can be resolved algebraically)	H-creative meta-diagrammatic abduction (as in creating a non-Euclidean geometry)

Table 1: 15 possible forms of abduction.

Based on my argument in Section 3 above that theoretical-model abduction is a precondition for any abduction since the goal of abduction is an explanation, and an explanation can be conceived as a theoretical model, it can be shown that this list of five things that can be abductively inferred is complete. If every abductively created insight presupposes theoretical-model abduction, then it must be possible to create also the elements abductively that occur in these models. Since these models are visible in the arguments that we must be able to construct according to my definition of an abductive insight, a complete list of things that can be abductively inferred must include every element of those arguments that is relevant from an explanatory point of view. Explanatory relevant in such an argument are facts, types (or theoretical concepts), and laws—and nothing else, I guess. In addition to these three, and model abduction itself, we need as a fifth form the one that allows the creation of the new representational systems since, as I argued also in Section 3, those systems are indispensable for creating theoretic models.

Georgia Institute of Technology

m.hoffmann@gatech.edu

References

Andersen, H. (2009). 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.

Anderson, D. (2005). Peirce and the Art of Reasoning. *Studies in Philosophy and Education, 24*(3-4), 277 - 289.

Bird, O. (1959). Peirce's Theory of Methodology. *Philosophy of science, 26*(3), 187-200.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: a mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36*(2), 421-441.

Blair, J. A. (2004). The Rhetoric of Visual Arguments. In C. A. Hill & M. Helmers (Eds.), *Defining Visual Rhetorics* (pp. 41-61): Lawrence Erlbaum Associates.

Boden, M. A. (2004 <1990>). *The creative mind. Myths and mechanisms* (2nd ed.). London ; New York: Routledge.

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.

Burton, R. G. (1999). A Neurocomputational Approach to Abduction. *Minds and Machines, 9*, 257–265.

Chasles, M. (1837). *Aperçu historique sur l'origine et le développement des methodes en géometrie*. Bruxelles (also Paris 1875).

Donohue, W., Kaufman, S., & Rogan, R. (Eds.). (forthcoming). *Framing in Negotiation: State of the Art*: Peter Lang.

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.

Field, J. V., & Gray, J. J. (1987). *The Geometrical Work of Girard Desargues*. New York Springer.

Friedman, M. (1974). Explanation and Scientific Understanding. *The Journal of Philosophy, 71*(1), 5-19.

Gabbay, D. M., & Woods, J. (2005). *The reach of abduction insight and trial*. Amsterdam Elsevier.

Hempel, C., & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science, 15*, 567-579.

Hilbert, D., & Cohn-Vossen, S. (1973 <1932>). *Anschauliche Geometrie*. Darmstadt: Wiss. Buchges.

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 (Knowledge development. A semiotic and pragmatic approach)*. Frankfurt am Main: Klostermann.

Hoffmann, M. H. G. (2005b). 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. (2010). Diagrams as Scaffolds for Creativity. *AAAI Workshops, North America*. Retrieved from http://aaai.org/ocs/index.php/WS/AAAIW10/paper/view/2027

Hoffmann, M. H. G. (in preparation). "... in a remote way Abduction rests upon diagrammatic reasoning ...".

Josephson, J. R., & Josephson, S. G. (Eds.). (1994). *Abductive Inference. Computation, Philosophy, Technology*. New York: Cambridge UP.

Kelle, U. (1994). *Empirisch begründete Theoriebildung: Zur Logik und Methodologie interpretativer Sozialforschung*. Weinheim: Dt. Studien-Verl.

Lewicki, R. J., Gray, B., & Elliott, M. (Eds.). (2003). *Making Sense of Intractable Environmental Conflicts. Concepts and Cases*. Washington - Covelo - London: Island Press.

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

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

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.

McKaughan, D. J. (2008). From Ugly Duckling to Swan: C. S. Peirce, Abduction, and the Pursuit of Scientific Theories. *Transactions of the Charles S. Peirce Society, 44*(3), 446-468.

Nesher, D. (2001). Peircean Epistemology of Learning and the Function of Abduction as the Logic of Discovery. *Transactions of the Charles S. Peirce Society, 37*(1), 23-59.

Norton, B. G. (2005). *Sustainability. A Philosophy of Adaptive Ecosystem Management*. Chicago University of Chicago Press.

Oevermann, U. (1991). Genetischer Strukturalismus und das sozialwissenschaftliche Problem der Erklärung der Entstehung des Neuen. In S. Müller-Doohm (Ed.), *Jenseits der Utopie: Theorie der Gegenwart* (pp. 352–433). Frankfurt am Main: Suhrkamp.

Otte, M. (1998). Limits of Contructivism: Kant, Piaget and Peirce. *Science & Education, 7*, 425–450.

Patokorpi, E. (2007). Logic of Sherlock Holmes in technology enhanced learning.

Educational Technology & Society, 10(1), 171-185.

Paul, G. (1993). Approaches to Abductive Reasoning: An Overview. *Artificial Intelligence Review, 7*, 109–152.

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.

Peirce (RLT). *Ch. S. Peirce, Reasoning and the Logic of Things*. Cambridge/London 1992: Harvard UP.

Peirce (SEM). *Charles S. Peirce, Semiotische Schriften* (Vol. I-III). Frankfurt a.M. 1986-1994: Suhrkamp.

Peirce, C. S. (1908). *Some Amazing Mazes: Explanation of Curiosity the First*: The Monist 18: 416-464; CP 4.594-638.

Prawat, R. S. (1999). Dewey, Peirce, and the Learning Paradox. *American Educational Research Journal, 36(1)*, 47–76.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, N.J.: Princeton University Press.

Schön, D. A., & Rein, M. (1994). *Frame reflection. Toward the resolution of intractable policy controversies*. New York: BasicBooks.

Schurz, G. (2008). Patterns of abduction. *Synthese, 164*(2), 201-234.

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

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1–36.

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 March 16, 2009).

Short, T. L. (1988). Hypostatic Abstraction in Empirical Science. *Grazer Philosophische Studien, 32*, 51–68.

Thagard, P. (2010). How Brains Make Mental Models. In L. Magnani (Ed.), *Model-Based Reasoning in Science & Technology* (pp. 447–461). Berlin Heidelberg: Springer-Verlag.

van Fraassen, B. C. (1980). *The scientific image*. Oxford: Clarendon Pr.

Wirth, U. (1999). *Diskursive Dummheit: Abduktion und Komik als Grenzphänomene des Verstehens*. Heidelberg: Winter.

Zeman, J. J. (1986). Peirce's Philosophy of Logic. *Transactions of the Charles S. Peirce Society*, *22*, 1–22.

---

[i] Peirce, LOS II 900 [1901]; cf. CP 5.182ff. [1903], 8.64 [1891].

[ii] Bechtel and Abrahamsen (2005) emphasize that "scientists who develop mechanistic explanations are not limited to linguistic representations and logical inference" since "they frequently employ diagrams to characterize mechanisms and simulations to reason about them" (p. 421). This, however, does not contradict my third condition since also "visual arguments" can be discussed as arguments (see Blair, 2004). Nevertheless, I would argue that even a diagrammatic explanation is acceptable as an explanation only if it is *possible* to provide a verbal argument. This follows, I think, directly from the definition of "mechanism" Bechtel and Abrahamsen provide: "A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena" (p. 423). Any "function" in this sense can be described as a conditional statement: if *p*, then *q*. In case it is not possible to describe a function as a conditional statement, then it simply is not a function.

[iii] See, for instance, Bonfantini & Proni, 1983; Eco, 1983; Shank & Cunningham, 1996; Magnani, 2001, 2009. In contrast to these pluralist approaches, McKaughan, 2008, argued recently that there is one—as yet neglected—interpretation that can be justified by a greater number of Peirce's remarks than the two following, famous alternatives: on the one hand, the interpretation that abduction is a process of generating new hypotheses and, on the other, that it is something like inference to the best explanation. According to what McKaughan calls the "Pursuitworthiness Interpretation," Peirce used the term "abduction" primarily for "systematic attempts to think about the qualities that factor into decisions about whether investigating an idea looks promising or

seems worthwhile. *Abductive reasoning makes practically grounded comparative recommendations about which available hypotheses are to be tested"* (p. 452; his emphasis). This, however, makes sense only if the goal is to find "the best explanation."

[iv] See, for example, Burton, 1999; Josephson & Josephson, 1994; Paul, 1993.

[v] See Anderson, 2005; Magnani, 2009; Nesher, 2001; Patokorpi, 2007; Prawat, 1999; Semetsky, 2005; Magnani, Nersessian, & Thagard, 1999; Magnani, 2001; Magnani & Nersessian, 2002; Shank & Cunningham, 1996.

[vi] For example in social sciences (Kelle, 1994; Oevermann, 1991) and literature: Wirth, 1999.

[vii] All italics in the following quotes are Schurz's.

[viii] See also Hoffmann, 2005b; Otte, 1998; Short, 1988; and Zeman, 1986. The meaning of "hypostatization"—formed from the Greek *hypostasis*—can be taken as "creating a new thing out of what is not a thing." Instead of using the Peircean concept of "hypostatic abstraction," recent learning theories in educational sciences are talking about "reification" (e.g. Sfard, 1991). Going back to a Latin root, the meaning of this term is exactly the same as that of the Greek root of "hypostatization."

[ix] See Magnani, 2001, who introduced the distinction between "selective" and "creative abduction."

[x] However, as Magnani, 2009, points out, in "inference to the best explanation" the evaluation of hypotheses is an essential component, which does not need to be the case in "selective abduction" (10).

[xi] The only texts I could find where Peirce discusses this concept clearly in the sense described above are Peirce, MS 754 [1907]: ISP 8 (wrongly counted), 6, and 7, and MS 318 [1907]: CSP 68 = ISP 225 = NEM III 491 and CSP 50 = ISP 42. The earlier usage of *theôrics* in 1902 (CP 1.278) is obviously not related to this later usage. As far as I know, the concept "theoric" has been discussed only by Bird, 1959, who characterizes everything that is creative in mathematics as "theoric" (194-195), neglecting thus the crucial difference between "theoric" and "theorematic" (see MS 754: ISP 8 and the two following notes). See also Hoffmann, 2005a, 170-186.

[xii] See CP 4.612-613. In this article C. S. Peirce, 1908), Peirce uses the term "theoric" 24 times (CP

4.602 Fn, 4.613-616, 619, 625, 627-629). In 4.613, the terms "theoric" and "theorematic" are clearly distinguished, in 4.627 less so.

[xiii] MS 318: CSP 50 = ISP 42 (the passage is also available in Helmut Pape's German translation in Peirce, SEM III 308-311; a connection to Desargues's theorem is also indicated in MS 754: ISP 8); "theóric deduction" is introduced here in contrast to "'corollarial' reasoning," a concept Peirce usually contrasts to "theorematic reasoning." Interestingly, the latter is just the term he uses to describe the proof of Desargues's theorem in NEM III 870-871 (1909). Without using any specific term, the proof is also discussed in Peirce, RLT 244-246 (1898), mentioned in NEM III 630 (1887) and EP II 174 (1903), and proved in NEM II 211-217 (c.1895) and in NEM III 846-847 (1909). The meanings of "theoric" and "theorematic" are in agreement when Peirce relates each of them in the passages mentioned above to the fact that "the demonstration of every considerable theorem of mathematics affords an instance of it" (MS 318) or "all the most theorems are of this nature" (NEM III 870). But besides the reference to *theorems*, the decisive difference between both is that while in theorematic reasoning "something else has to be added" to "the Diagram of the truth of the Premisses" (NEM III), the "theóric deduction" consists exclusively in "looking at facts from a novel point of view" (MS 318)—without adding something. In MS 754: ISP 8 (1907), Peirce tries to clarify the terms theoric, theorematic, and corrolarial in the following notes: "I formerly, quite dubiously, divided Deductions into the Corollarial & the Theorematic. Explain these. Deduction will better be called Demonstration. But further study leads me to lop off a corollarial part from the Theorematic Deductions, which follows that part that originates a new point of view. This part of the theorematic procedure, I will call theôric reasoning. It is very plainly allied to retroduction, from which it only differs as far as I now see in being indisputable." (I am thankful to André de Tienne from the Peirce Edition Project for transcribing the page ISP 8 for me, and for confirming its correct location between ISP 5 and 6.)

[xiv] Here for the first time written without an apostrophe over the "o" and without double quotation marks as two pages earlier in the manuscript (see the previous note). Oddly enough, on this previous page Peirce writes just above and on the right of "theóric" the word "theo'ric"—both exactly in this form, including the quotations marks. In contrast to this, we find "theō'ric transformation" (in exactly this form, but without quotation marks) in the same MS 318 on page CSP 68 = ISP 225 = NEM III 491. This form is closer to the Greek spelling of *theoria*, because "ō" is usually chosen as the Latin transliteration of the Greek "ω." See also MS 754: ISP 8, 6, and 7 (1907), where Peirce writes "theôric" (without quotation marks). From a language development point of view we would expect that more complicated forms precede simpler ones like "theoric."

[xv] It should be mentioned that this "theoric step" must be ascribed to Desargues already (Field & Gray, 1987, 160). Karl Georg Christian von Staudt formulated the first complete theory of projective geometry that allowed dealing with all the special cases of Desargues's theorem (e.g., the case when the two planes are parallel). Peirce could have known this fact from Chasles, 1837, whom he mentions several times (CP 3.555; NEM III 103-104 = RLT 244-245 and NEM III 1018-9).

[xvi] See, for example, Schön & Rein, 1994, Lewicki, Gray, & Elliott, 2003 and Donohue, Kaufman, & Rogan, forthcoming.