Logic Journal of IGPL Advance Access published April 4, 2013

Is abduction ignorance-preserving? Conventions, models and fictions in science

LORENZO MAGNANI*, *Department of Humanities, Philosophy Section, and Computational Philosophy Laboratory, University of Pavia, Pavia, 27100, Italy*

> Of the three Universes of Experience familiar to us all, the first comprises all mere Ideas, those airy nothings to which the mind of poet, pure mathematician, or another might give local habitation and a name within that mind. Their very airy-nothingness, the fact that their Being consists in mere capability of getting thought, not in anybody's Actually thinking them, saves their Reality.
>
> Charles Sanders Peirce, *A Neglected Argument for the Reality of God*, 1908

Abstract

Abduction is a procedure in which something that lacks classical explanatory epistemic virtue can be accepted because it has virtue of another kind: Gabbay and Woods (2005, *The Reach of Abduction*) contend (GW-model) that abduction presents an *ignorance-preserving* or (ignorance-mitigating) character. From this perspective abductive reasoning is a *response* to an ignorance-problem; through abduction the basic ignorance—that does not have to be considered a total 'ignorance'—is neither solved nor left intact. Abductive reasoning is an ignorance-preserving accommodation of the problem at hand. Is abduction really ignorance-preserving? To better answer this question I will take advantage of my *eco-cognitive model* (EC-model) of abduction and of three examples taken from the areas of both philosophy and epistemology. It will be illustrated that through abduction, knowledge can be enhanced, even when abduction is not considered an inference to the best explanation (IBE) in the classical sense of the expression, i.e. an inference necessarily characterized by an empirical evaluation phase, or an inductive phase, as Peirce called it. (1) Peirce provides various justifications of the knowledge enhancing role of abduction, even when abduction is not considered an IBE in the classical sense of the expression, i.e. an inference necessarily characterized by an empirical evaluation phase, or inductive phase. These justifications basically resort to the conceptual exploitation of *evolutionary* and *metaphysical* ideas, which clearly show that abduction is constitutively akin to truth, even if certainly always ignorance-preserving or mitigating in the sense that the 'absolute truth' is never reached through abduction; (2) in empirical science abducing *conventions* favours and increases knowledge even if these hypotheses remain evidentially inert—at least in the sense that it is not possible to empirically falsify them. Consequently abduced conventions are evidentially inert but knowledge enhancing at the rational level of science; (3) in science we do not have to confuse the process of abducing models with the process of abducing fictions: the recent epistemological conundrum concerning fictionalism presents to us the epistemic situation in which the models abduced by scientists reveal themselves not to be 'airy nothings' at all, and certainly different in their gnoseological status from literary fictions. Scientific models instead play fundamental 'rational' knowledge enhancing roles: in a static perspective (e.g. when inserted in a textbook) scientific models can appear fictional to the epistemologist, but their fictional character disappears if a dynamic perspective is adopted. Abduction in scientific model-based reasoning is not a suspicious process of guessing fictions.

Keywords: Abduction, ignorance-preservation, models, fictions, conventions.

*E-mail: lmagnani@unipv.it

doi:10.1093/jigpal/jzt012

1 Is abduction ignorance-preserving?

1.1 The ignorance-preserving character of abduction

As I have illustrated in my book on abductive cognition, [32, ch. 2] following Gabbay and Woods' contention, it is clear that '[…] abduction is a procedure in which something that lacks epistemic virtue is accepted because it has virtue of another kind' [15, p. 62]. For example: 'Let S be the standard that you are not able to meet (e.g. that of mathematical proof). It is possible that there is a lesser epistemic standard S' (e.g. having reason to believe) that you do meet' [61, ch. 10]. Focusing attention on this cognitive aspect of abduction, and adopting a logical framework centred on practical agents, [15] contend that abduction (basically seen as a *scant-resource* strategy, which proceeds in absence of knowledge) presents an *ignorance-preserving* (or, better, an *ignorance-mitigating*) character. Of course '[…] it is not at all necessary, or frequent, that the abducer be wholly in the dark, that his ignorance be total. It needs not be the case, and typically isn't, that the abducer's choice of a hypothesis is a blind guess, or that nothing positive can be said of it beyond the role it plays in the subjunctive attainment of the abducer's original target (although sometimes this is precisely so)' (cit.). In this perspective, abductive reasoning is a *response* to an ignorance-problem: one has an ignorance-problem when one has a cognitive target that cannot be attained on the basis of what one currently knows. Ignorance problems trigger one or other of three responses. In the first case, one overcomes one's ignorance by attaining some additional knowledge (subduance). In the second instance, one yields to one's ignorance (at least for the time being) (surrender). In the third instance, one abduces [61, ch. 10] and so has some positive basis for new action even if in the presence of the constitutive ignorance.

From this perspective the general form of an abductive inference can be symbolically rendered as follows. Let α be a proposition with respect to which you have an ignorance-problem. Putting T for the agent's epistemic target with respect to the proposition α at any given time, K for his knowledge-base at that time, K^* for an immediate accessible successor-base of K that lies within the agent's means to produce in a timely way,[1] R as the attainment relation for T, $\rightsquigarrow$ as the *subjunctive* conditional relation, H as the agent's hypothesis, $K(H)$ as the revision of K upon the addition of H, $C(H)$ denotes the conjecture of H and H^c its activation. The general structure of abduction can be illustrated as follows (GW-schema):[2]

1. $T!\alpha$	[setting of T as an epistemic target with respect to a proposition α]
2. $\neg(R(K, T)$	[fact]
3. $\neg(R(K^*, T)$	[fact]
4. $H \notin K$	[fact]
5. $H \notin K^*$	[fact]
6. $\neg R(H, T)$	[fact]
7. $\neg R(K(H), T)$	[fact]
8. If $H \rightsquigarrow R(K(H), T)$	[fact]

[1] K^* is an accessible successor of K to the degree that an agent has the know-how to construct it in a timely way; i.e. in ways that are of service in the attainment of targets linked to K. For example, if I want to know how to spell 'accommodate', and have forgotten, then my target cannot be hit on the basis of K, what I now know. But I might go to my study and consult the dictionary. This is K^*. It solves a problem originally linked to K.

[2] That is Gabbay and Woods schema.

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

It is easy to see that the distinctive epistemic feature of abduction is captured by the schema. It is a given that H is not in the agent's knowledge-set. Nor is it in its immediate successor. Since H is not in K, then the revision of K by H is not a knowledge-successor set to K. Even so, $H \rightsquigarrow (K(H), T)$. So we have an ignorance-preservation, as required (cf. [61, ch. 10]).

[*Note*: Basically, line 9. indicates that H has no more plausible or relevant rival constituting a greater degree of subjunctive attainment. Characterizing the S_i is the most difficult problem for abductive cognition, given the fact that in general there are many possible candidate hypotheses. It involves for instance the *consistency* and *minimality* constraints.[3] These constraints correspond to the lines 4 and 5 of the standard AKM-schema of abduction,[4] which is illustrated as follows:

1. E
2. $K \not\vdash E$
3. $H \not\vdash E$
4. $K(H)$ is consistent
5. $K(H)$ is minimal
6. $K(H) \vdash\!\!\!\!\sim E$
7. Therefore, H.

[15, pp. 48–49].

where of course the conclusion operator $\vdash\!\!\!\!\sim$ cannot be classically interpreted].[5]

Finally, in the GW-schema $C(H)$ is read 'It is justified (or reasonable) to conjecture that H' and H^c is its activation, as the basis for *planned* 'actions'.

In sum, in the GW-schema T cannot be attained on the basis of K. Neither can it be attained on the basis of any successor K^* of K that the agent knows then and there how to construct. H is not in K: H is a hypothesis that when reconciled to K produces an updated $K(H)$. H is such that if it were true, then $K(H)$ would attain T. The problem is that H is *only hypothesized*, so that the truth is not assured. Accordingly Gabbay and Woods contend that $K(H)$ *presumptively* attains T. That is, having hypothesized that H, the agent just 'presumes' that his target is now attained. Given the fact that presumptive attainment is not attainment, the agent's abduction must be considered as preserving the ignorance that already gave rise to her (or its, in the case for example of a machine) initial ignorance-problem. Accordingly, abduction does not have to be considered the 'solution' of

[3] I have shown [32, ch. 2, subsection 2.3.1] that, in the case of inner processes in organic agents, this sub-process—here explicitly modelled thanks to a formal schema—is considerably implicit, and so also linked to unconscious ways of inferring, or even, in Peircean terms, to the activity of the instinct [41, 8.223] and of what Galileo called the *lume naturale* [41, 6.477], i.e. the innate fair for guessing right I will decribe below in subsection 1.4. This and other cognitive aspects can be better illustrated thanks to the alternative EC-model of abduction I will sketch below, subsection 1.2.

[4] The classical schematic representation of abduction is expressed by what [15] call AKM-schema, which is contrasted to their own (GW-schema), which I am just explaining in this subsection. For A they refer to Aliseda [1, 2], for K to Kowalski [24], Kuipers [25] and Kakas *et al.* [20], for M to Magnani [30] and Meheus [36]. A detailed illustration of the AKM-schema is given in [32, ch. 2, subsection 2.1.3].

[5] The target has to be an explanation and $K(H)$ bears R^{pres} [i.e. the relation of presumptive attainment] to T only if there is a proposition V and a consequence relation $\vdash\!\!\!\!\sim$ such that $K(H) \vdash\!\!\!\!\sim V$, where V represents a *payoff proposition* for T. In turn, in this schema explanations are interpreted in consequentialist terms. If E is an explanans and E' an explanandum the first explains the second only if (some authors further contend if and only if) the first implies the second. It is obvious to add that the AKM-schema embeds a D-N (deductive-nomological) interpretation of explanation, as I have already stressed in [30, p. 39].

an ignorance problem, but rather a response to it, in which the agent reaches presumptive attainment rather than actual attainment. $C(H)$ expresses the conclusion that it follows from the facts of the schema that H is a worthy object of conjecture. It is important to note that in order to solve a problem it is not necessary that an agent actually conjectures a hypothesis, but it is necessary that she states that the hypothesis is *worthy of conjecture*.

It is remarkable that in the above schema.

> […] $R(K(H), T)$ is false and yet that $H \rightsquigarrow (K(H), T)$ is true. Let us examine a case. Suppose that your target T is to know whether α is true. Suppose that, given your present resources, you are unable to attain that target. In other words, neither your K nor your K^* enables you to meet your target. Let H be another proposition that you don't know. So $K(H)$ is not a knowledge-set for you. On the principle that you can't get to know whether α on the basis of what you don't know, $K(H)$ won't enable you to attain T either. This is a point of some subtlety. Pages ago, weren't we insisting that there are contexts—autoepistemic contexts—in which not knowing something is a way of getting to know something else? No, we said that not knowing something was a way of getting to presume something else. But just to be clear, let us point out that in the GW-schema α and H are not candidates for the autoepistemic inference of α from H or $K(H)$. So $R(K(H), T)$ is false. $H \rightsquigarrow (K(H), T)$ is different. It says, subjunctively, that if H were true, then the result of adding H to K would attain T. Clearly this can be true while, for the same H, K and T, $R(K(H), T)$ is false [61, ch. 8].

Finally, considering H justified to conjecture is not equivalent to considering it justified to accept/activate it and eventually to send H to experimental trial. H^c denotes the *decision* to release H for further premissory work in the domain of enquiry in which the original ignorance-problem arose, i.e. the activation of H as a positive *cognitive* basis for action. Woods usefully observes:

> There are lots of cases in which abduction stops at line 10, that is with the conjecture of the hypothesis in question but not its activation. When this happens, the reasoning that generates the conjecture does not constitute a positive basis for new action, that is for acting *on* that hypothesis. Call these abductions *partial* as opposed to full. Peirce has drawn our attention to an important subclass of partial abductions. These are cases in which the conjecture of H is followed by a decision to submit it to experimental test. Now, to be sure, doing this is an action. It is an action *involving* H but it is not a case of acting *on* it. In a full abduction, H is activated by being released for inferential work in the domain of enquiry within which the ignorance-problem arose in the first place. In the Peircean cases, what counts is that H is withheld from such work. Of course, if H goes on to test favourably, it may then be released for subsequent inferential engagement [57].

We have to remember that this process of evaluation and so of activation of the hypothesis, is not abductive, but inductive, as Peirce contended. Woods adds: 'Now it is quite true that epistemologists of a certain risk-averse bent might be drawn to the admonition that partial abduction is as good as abduction ever gets and that complete abduction, inference-activation and all, is a mistake that leaves any action prompted by it without an adequate rational grounding. This is not an unserious objection, but I have no time to give it its due here. Suffice it to say that there are real-life contexts of reasoning in which such conservatism is given short shrift, in fact is ignored altogether. One of these contexts is the criminal trial at common law' [57].

In the framework of the GW-schema it cannot be said that testability is intrinsic to abduction, such as it is instead maintained in the case of some passages of Peirce's writings.[6] This activity of testing, I repeat, which in turn involves degrees of risk proportioned to the strength of the conjecture, is strictly cognitive/epistemic and inductive in itself, e.g. an experimental test, and it is an intermediate step to release the abduced hypothesis for inferential work in the domain of enquiry within which the ignorance-problem arose in the first place.

Through abduction the basic ignorance—that does not have to be considered total 'ignorance'—is neither solved nor left intact: it is an ignorance-preserving accommodation of the problem at hand, which 'mitigates' the initial cognitive 'irritation' (Peirce says 'the irritation of doubt').[7] As I have already stressed, further action can be triggered—in a defeasible way—either to find further abductions or to 'solve' the ignorance problem, possibly leading to what the 'received view' has called the *inference to the best explanation* (IBE).

It is clear that in the framework of the GW-schema the IBE—if considered as a truth conferring achievement justified by the empirical approval—cannot be a case of abduction, because abductive inference is constitutively ignorance-preserving. In this perspective the IBE involves the generalizing[8] and evaluating role of *induction*. Of course it can be said that the requests of originary thinking are related to the depth of the abducer's ignorance.

1.2 The EC-model of abduction: cutdown and fill-up problems

From a general philosophical perspective (with, and beyond, Peirce) the condition 9. (cf. the GW-schema) is, as Woods himself admits 'more a hand-wave than a real condition. Of course the devil is in the details. […] I myself I am not sure' [58, p. 242]. Obviously consistency and minimality constraints were emphasized in the 'received view' on abduction established by many classical logical accounts, more oriented to illustrate selective abduction [30]—e.g. in diagnostic reasoning, where abduction is merely seen as an activity of 'selecting' from an encyclopedia of pre-stored hypotheses—than to analyse *creative* abduction (abduction that generates new hypotheses).[9]

For example, to stress the puzzling status of the consistency requirement, it is here sufficient to note that Paul Feyerabend, in *Against Method* [13], correctly attributes a great importance to the role of contradiction in generating hypotheses, also against the role of similarity, and so implicity celebrates the value of creative abductive cognition. Speaking of induction and not of abduction (this concept was relatively unknown at the level of the international philosophical community at that time), he establishes a new 'counterrule'. This is the opposite of the neopositivistic one that it is 'experience' (or 'experimental results') which constitutes the most important part of our scientific

[6] When abduction stops at line 10. (cf. the GW-schema), the agent is not prepared to accept $K(H)$, because of supposed adverse consequences.

[7] 'The action of thought is excited by the irritation of doubt, and ceases when belief is attained; so that the production of belief is the sole function of thought' [43, p. 261].

[8] By illustrating abductive/inductive reasoning of preservice elementary majors on patterns that consist of figural and numerical cues in learning elementary mathematics Rivera and Becker monitor the subsequent role of induction. In performing the abductive task to the general form/hypothesis the subjects referred to the fact they immediately saw a relationship among the drawn cues in terms of relational similarity '[…] within classes in which the focus was *not* on the individual clues in a class *per se* but on a possible invariant relational structure that was perceived between and, thus, projected onto the cues' [47, p. 151]. Through the follow-up inductive stage of generalizations the subjects tested the hypotheses just examining *extensions* (new particular cases beyond what was available at the beginning of the reasoning process). This process was also able to show subjects's disconfirmation capacities: they acknowledged their mistakes in generating a bad induction, which had to be abandoned, in so far as they were checked as insufficient in fully capturing in symbolic terms a general attribute that would yield the total number of toothpicks in new generated cues.

[9] I have proposed the dichotomic distinction between selective and creative abduction in [30].

empirical theories, a rule that formed the core of the so-called 'received view' in philosophy of science (where inductive generalization, confirmation, and corroboration play a central role). The counterrule '[…] advises us to introduce and elaborate hypotheses which are inconsistent with well-established theories and/or well-established facts. It advises us to proceed counterinductively' [13, p. 20]. Counterinduction is seen more reasonable than induction, because appropriate to the needs of creative reasoning in science: '[…] we need a dream-world in order to discover the features of the real world we think we inhabit' (p. 29). We know that counterinduction, i.e. the act of introducing, inventing and generating new inconsistencies and anomalies, together with new points of view incommensurable with the old ones, is congruous with the aim of inventing 'alternatives' (Feyerabend contends that 'proliferation of theories is beneficial for science'), and very important in all kinds of creative reasoning.

Since for many abduction problems there are—usually—many guessed hypotheses, the abducer needs reduce this space to one; this means that the abducer has to produce the best choice among the members of the available group: 'It is extremely difficult to see how this is done, both formally and empirically. Clause (9) [in the GW-model] is a place-holder for two problems, not one. There is the problem of finding criteria for hypothesis *selection*. But there is the prior problem of specifying the conditions for *thinking up* possible candidates for selection. The first is a 'cutdown' problem. The second is a 'fill-up problem'; and with the latter comes the received view that it is not a problem for logic' ([58, p. 243] emphasis added).

Here we touch the core of the ambiguity of the ignorance-preserving character of abduction. Why?

- Because the cognitive processes of generation (fill-up) and of selection (cutdown) can both be sufficient—even in absence of the standard inductive evaluation phase—to *activate* and accept [clause (11) of the GW-schema above] an abductive hypothesis, and so to reach cognitive results relevant to the context (often endowed with a knowledge enhancing outcome, as I will illustrate below in the subsections 2.1 and 2.2). In these cases the instrumental aspects[10] (which simply enable one's target to be hit) often favour both abductive generation and abductive choice, and they are not necessarily intertwined with plausibilistic concerns, such as consistency and minimality.

In these special cases the best choice is immediately reached without the help of an experimental trial (which fundamentally characterizes the received view of abduction in terms of the so-called 'inference to the best explanation'). Not only, we have to strongly note that the generation process alone can suffice, like it is demonstrated by the case of human *perception*, where the hypothesis generated is immediate and unique. Indeed, perception is considered by Peirce, as an 'abductive' fast and uncontrolled (and so automatic) knowledge-production procedure. Perception, in this philosophical perspective, is a vehicle for the instantaneous retrieval of knowledge that was previously structured in our mind through more structured inferential processes. Peirce says: 'Abductive inference shades into perceptual judgment without any sharp line of demarcation between them' [42, p. 304]. By perception, knowledge constructions are so instantly reorganized that they become habitual and diffuse and do not need any further testing: '[…] a fully accepted, simple and interesting inference tends to obliterate all recognition of the uninteresting and complex premises from which it was derived' [41, 7.37].[11]

My abrupt reference to perception as a case of abduction (in this case I strictly follow Peirce) does not have to surprise the reader. Indeed, at the of centre of my perspective on cognition is the emphasis

[10] Cf. above subsection 2.1.

[11] A relatively recent cognitive research related to artificial intelligence presents a formal theory of robot perception as a form of abduction, so reclaiming the rational relevance of the speculative anticipation furnished by Peirce, cf. [49].

on the 'practical agent', of the individual agent operating 'on the ground', i.e. in the circumstances of real life. In all its contexts, from the most abstractly logical and mathematical to the most roughly empirical, I always emphasize the cognitive nature of abduction. Reasoning is something performed by cognitive systems. At a certain level of abstraction and as a first approximation, a cognitive system is a triple (A, T, R), in which A is an *agent*, T is a *cognitive target* of the agent, and R relates to the *cognitive resources* on which the agent can count in the course of trying to meet the target-information, time and computational capacity, to name the three most important. My agents are also *embodied distributed cognitive systems*: cognition is embodied and the interactions between brains, bodies and external environment are its central aspects. Cognition is occurring taking advantage of a constant exchange of information in a complex distributed system that crosses the boundary between humans, artefacts and the surrounding environment, where also instinctual and unconscious abilities play an important role. This interplay is especially manifest and clear in various aspects of abductive cognition.[12]

It is in this perspective that we can appropriately consider perceptual abduction as a fast and uncontrolled knowledge production, that operates for the most part automatically and out of sight, so to speak. This means that—at least in this light—GW-schema is not canonical for abduction. The schema illustrates what I call 'sentential abduction' [32, ch. 1], i.e. abduction rendered by symbols carrying propositional content. It is hard to encompass in this model cases of abductive cognition such as perception or the generation of models in scientific discovery (cf. below subsection 2.2). My perspective adopts the wide Peircean philosophical framework, which approaches 'inference' *semiotically* (and not simply '*logically*'): Peirce distinctly says that all inference is a form of sign activity, where the word sign includes 'feeling, image, conception, and other representation' [41, 5.283]. It is clear that this semiotic view is considerably compatible with my perspective on cognitive systems as embodied and distributed systems: the GW-schema is instead only devoted to illustrate, even if in a very efficacious way, a subset of the cognitive systems abductive activities, the ones that are performed taking advantage of explicit propositional contents. Woods seems to share this conclusion: '[…] the GW-model helps get us started in thinking about abduction, but it is nowhere close, at any level of abstraction, to running the whole show. It does a good job in modelling the ignorance-preserving character of abduction; but, since it leaves the S_i of the schema's clause (T) unspecified, it makes little contribution to the fill-up problem' [58, p. 244].

In a wide eco-cognitive perspective the cutdown and fill-up problems in abductive cognition appear to be spectacularly *contextual*.[13] I lack the space to give this issue appropriate explanation but it suffices for the purpose of this study to remember that, e.g. one thing is to abduce a model or a concept at the various levels of scientific cognitive activities, where the aim of reaching rational knowledge dominates, another thing is to abduce a hypothesis in literature (e.g. a fictional character), or in moral reasoning (the adoption/acceptation of a hypothetical judgment as a trigger for moral actions). However, in all these cases abductive hypotheses which are evidentially inert are accepted and activated as a basis for action, even if of different kind.

The backbone of this approach can be found in the manifesto of my EC-model of abduction in [32]. It might seem awkward to speak of 'abduction of a hypothesis in literature,' but one of the fascinating aspects of abduction is that not only it can warrant for scientific discovery, but

[12] It is interesting to note that recent research on Model Checking in the area of AST (Automated Software Testing) takes advantage of this eco-cognitive perspective, involving the manipulative character of model-based abduction in the practice of adapting, abstracting and refining models that do not provide successful predictions. Cf. [3], this volume.

[13] Some acknowledgement of the general contextual character of these kinds of criteria, and a good illustration of the role of coherence, unification, explanatory depth, simplicity and empirical adequacy in the current literature on scientific abductive best explanation, is given in [29].

for other kinds of creativity as well. We must not necessarily see abduction as a *problem solving device* that sets off in response to a cognitive irritation/doubt: conversely, it could be supposed that esthetic abductions (referring to creativity in art, literature, music, etc.) arise in response to some kind of esthetic irritation that the author (sometimes a *genius*) perceives in herself or in the public. Furthermore, not only esthetic abductions are free from empirical constraints in order to become the 'best' choice: as I am showing throughout this article, many forms of abductive hypotheses in traditionally perceived-as-rational domains (such as the setting of initial conditions, or axioms, in physics or mathematics) are relatively free from the need of an empirical assessment. The same could be said of moral judgment: they are eco-cognitive abductions, inferred upon a range of internal and external cues and, as soon as the judgment hypothesis has been abduced, it immediately becomes prescriptive and 'true,' informing the agent's behaviour as such. Assessing that there is a common ground in all of these works of what could be broadly defined as 'creativity' does not imply that all of these forms of creativity are the same, contrarily it should spark the need for firm and sensible categorization: otherwise it would be like saying that to construct a doll, a machine-gun and a nuclear reactor are all the same thing because we use our hands in order to do so!

To conclude this subsection I have to say some words about the role of heuristics. From an eco-cognitive point of view, in more hybrid and multimodal (cf. [32, ch. 4]) (not merely inner) abductive processes, such as in the case of manipulative abduction,[14] the *assessment/acceptation* of a hypothesis is reached—and constrained—taking advantage of the gradual acquisition of consecutive external information with respect to future interrogation and control, and not necessarily thanks to a final and actual experimental test, in the classical sense of empirical science.

Hintikka implicitly acknowledges the multimodality and hybridity of what I call *selective abduction* when, taking advantage of the intellectual atmosphere of his Socratic interrogative epistemology, observes that '[…] abduction as a method of guessing is based on the variety of different possible sources of answers. Such 'informants' must include not only testimony, observation and experiments, but the inquirer's memory and background knowledge' [18, p. 56]. Moreover, Hintikka further notes that also 'creative abduction', generated by a kind of *oracle*, is often needed: 'But what can an inquirer do when all such sources fail to provide an answer to a question? Obviously the best the inquirer can do is make an informed guess. For the purposes of a general theory of inquiry, what Peirce calls 'intelligent guessing' must therefore be recognized as one of the many possible 'oracles', alias sources of answers. Peirce may very well have been more realistic than I have so far been in emphasizing the importance of this particular 'oracle' in actual human inquiry' (*ibid*.).

In summary, at least four kinds of actions can be involved in the manipulative abductive processes (and we would have to also take into account the motoric aspect (1) of inner 'thoughts' too). In the eco-cognitive interplay of abduction the cognitive agent further triggers internal *thoughts* 'while' modifying the environment and so (2) acting on it (thinking through doing). In this case the 'motor actions' directed to the environment have to be intended as part and parcel of the whole embodied abductive inference, and so have to be distinguished from the *final* (3) 'actions' as a possible consequence of the reached abductive result.

[14] The concept of *manipulative abduction*—which also takes into account the external dimension of abductive reasoning in an eco-cognitive perspective—captures a large part of scientific thinking where the role of action and of external models (e.g. diagrams) and devices is central, and where the features of this action are implicit and hard to be elicited. Action can provide otherwise unavailable information that enables the agent to solve problems by starting and by performing a suitable abductive process of generation and/or selection of hypotheses. Manipulative abduction happens when we are thinking through doing and not only, in a pragmatic sense, about doing (cf. [32, ch. 1]). Cf. also below, subsection 2.2.2.

In this perspective the proper experimental test involved in the Peircean evaluation phase, which for many researchers reflects in the most acceptable way the idea of abduction as IBE, just constitutes a *special* subclass of the process of the adoption of the abductive hypothesis—the one which involves a terminal kind (4) of actions (experimental tests), and should be considered ancillary to the nature of abductive cognition, and inductive in its essence. We have indeed to remark again that in Peirce's mature perspective on abduction as embedded in a cycle of reasoning, induction just plays an evaluative role. Hintikka usefully notes, and I agree with him, that Peirce was right in denying the role of 'naked' induction in forming new hypotheses:

> Many philosophers would probably bracket abductive inference with inductive inference. Some would even think of all ampliative inference as being, at bottom, inductive. In this matter, however, Peirce is one hundred percent right in denying the role of naked induction in forming new hypotheses. […] It might seem that the critical and evaluative aspect of inquiry that Peirce called inductive still remains essentially different from the deductive and abductive aspects. A common way of thinking equates all ampliative inferences with inductive ones. Peirce was right in challenging this dichotomy. Rightly understood, the ampliative versus non-ampliative contrast becomes a distinction between interrogative (ampliative) and deductive steps of argument. As in Peirce, we also need over and above these two also the kind of reasoning that is involved in testing the propositions obtained as answers to questions. I do not think that it is instructive to call such reasoning inductive, but this is a merely terminological matter [18, pp. 52 and 55].

In absence of empirical evaluation, can we attribute the *pure* abductive inclination to produce right guesses indicated by Peirce, conductive to the acquisition of truth, to the *reliability* of the process? Yes, we can, but only if we take into account the following warning, still illustrated by Hintikka: 'Many contemporary philosophers will assimilate this kind of justification to what is called a reliabilist one. Such reliabilist views are said to go back to Frank Ramsey, who said that 'a belief was knowledge if it is (1) true, (2) certain, (3) *obtained by a reliable process*' (emphasis added). Unfortunately for reliabilists, such characterizations are subject to the ambiguity that was pointed out earlier. By a reliable process one can mean either a process in which each step is conducive to acquiring and/or maintaining truth or closeness to truth, or one that as a whole is apt to lead the inquirer to truth. Unfortunately, most reliabilists unerringly choose the wrong interpretation—namely, the first one. As was pointed out earlier, the true justification of a rule of abductive inference is a strategic one' [18, p. 57]. The important thing is to stress that this strategic justification *does not warrant* any specific step of the whole process. Let us remember that abduction certainly provides new information into an argument, but this is not necessarily a true information, because it is not implied by what it is already known or accepted but it is constitutively hypothetical—i.e. ignorance-preservation is constitutive, from the general logico-philosophical point of view, and Hintikka is in tune with this assumption.

1.3 Abductive virtues vindicated. How does abduction supply knowledge?

Even if abduction, in the perspective of the formal GW-model above, is ignorance-preserving (or ignorance-mitigating), truth can easily emerge: we have to remember that Peirce sometimes contended that abduction 'come to us as a flash. It is an act of insight' [41, 5.181] but nevertheless possesses a mysterious power of 'guessing right' [41, 6.530]. Consequently abduction, preserves ignorance, in the logical sense I have illustrated above, but also can provide truth because has the

power of guessing *right*. We have also contended that in the logical framework above the IBE—if considered as a truth conferring achievement justified by empirical approval—cannot be a case of abduction, because abductive inference is instead constitutively ignorance-preserving.

If we say that truth can be reached through a 'simple' abduction (not intended as involving an evaluation phase, i.e. coinciding with the whole IBE, fortified by an empirical evaluation), it seems we confront a manifest incoherence. In this perspective it is contended that even simple abduction can provide truth, even if it is epistemically 'inert' from the empirical perspective. Why? We can solve the incoherence by observing that we should be compelled to consider abduction as ignorance-preserving only if we consider the empirical test *the only way* of conferring truth to a hypothetical knowledge content. This clause being accepted, in the framework of the technical logical model of abduction I have introduced above the ignorance-preservation appears natural and unquestionable. However, if we admit that there are ways to accept a hypothetical knowledge content different from the empirical test, simple abduction is not necessarily constitutively ignorance-preserving: in the end we are dealing with a disagreement about the nature of *knowledge*, as Woods himself contends. As I have indicated at the end of the previous subsection, those who consider abduction as an IBE—i.e. as a truth conferring achievement involving empirical evaluation—obviously cannot consider abductive inference as ignorance-preserving. Those who consider abduction as a mere activity of guessing are more inclined to accept its ignorance-preserving character.

However, *we are objecting that abduction is in this last case still knowledge enhancing*.

At this point two important consequences concerning the meaning of the word *ignorance* in this context have to be illustrated:

1. abduction, also when intended as an IBE in the 'classical' sense I have indicated above, is always *ignorance-preserving* because abduction represents a kind of reasoning that is constitutively provisional, and you can withdraw previous abductive results (even if empirically confirmed, i.e. appropriately considered 'best explanations'), in presence of new information. From the logical point of view this means that abduction represents a kind of non-monotonic reasoning, and in this perspective we can even say that abduction interprets the 'spirit' of modern science, where truths are never stable and absolute. Peirce also emphasized the 'marvelous self-correcting property of reason' in general [41, 5.579]. So to say, abduction incarnates the human perennial search of new truths and the human Socratic awareness of a basic ignorance which can only be attenuated/mitigated. In sum, in this perspective abduction always preserves ignorance because it reminds us we can reach truths that can always be withdrawn; ignorance removal is at the same time constitutively related to ignorance regaining;
2. even if ignorance is preserved in the sense I have just indicated, which coincides with the spirit of modern science, abduction is also knowledge enhancing because new truths can be and 'are' discovered which *are not necessarily best explanations intended as hypotheses which are empirically tested.*

A similar argumentation, which resorts to better explain the conundrum of abduction as ignorance-preserving, is provided by Woods, who notes that some philosophers accept the Gabbay–Woods schema (GW-schema) for abduction but at the same time dislike its commitment to the ignorance-preservation claim. Woods' answer resorts to say that this hesitancy flows from how those philosophers *epistemologically* approach the general question of knowledge. It is not logic of abduction in question but the epistemological adopted perspective [61, ch. 10]. I have just said that knowledge can be attained in the absence of evidence; there are propositions about the world which turn to be true by virtue of considerations that lend them no evidential/empirical weight. They are true beliefs

that are not justified on the basis of evidence. Is abduction related to the generation of knowledge contents of this kind? Yes it is.

Abduction is guessing reliable hypotheses, and humans are very good at it; abduction is akin to truth: it is especially in the case of empirical scientific cognition that abduction reveals its more representative epistemic virtues, because it provides hypotheses, models, ideas, thoughts experiments, etc., which, even if *devoid of initial* evidential support, constitute the fundamental rational building blocks for the generation of new laws and theories which only later on will be solidly empirically tested.

In the following sections of this study I aim at illustrating this intrinsic character of abduction, which shows why we certainly can logically consider it a kind of ignorance-preserving cognition, but at the same time a cognitive process that can enhance knowledge at various level of human cognitive activities, even if the empirical evaluation lacks.

1.4 Why does abduction enhance knowledge? Instinct, inference and synechism: Mind and matter intertwined

Peirce provides various justifications of the productive gnoseological role of abduction. They basically resort to the conceptual exploitation of evolutionary and metaphysical ideas, which clearly show that abduction is constitutively akin to truth, certainly ignorance-preserving—because the 'absolute truth' is never reached through abduction—but also knowledge enhancing. Peirce himself notes that abductive guesses are belief-inducing and truth-making. Not only, it cannot be said that unevidenced belief is itself evidence of malfunction and disorder, and so of falsity.

First of all Peirce considers hypothesis generation a largely instinctual endowment[15] of human beings given by God or related to a kind of Galilean '*lume naturale*': 'It is a primary *hypothesis* underlying all abduction that the human mind is akin to the truth in the sense that in a finite number of guesses it will light upon the correct hypothesis' [41, 7.220]. Again, the example of the innate ideas of 'every little chicken' is of help to describe this human instinctual endowment:

> How was it that man was ever led to entertain that true theory? You cannot say that it happened by chance, because the possible theories, if not strictly innumerable, at any rate exceed a trillion – or the third power of a million; and therefore the chances are too overwhelmingly against the single true theory in the twenty or thirty thousand years during which man has been a thinking animal, ever having come into any man's head. Besides, you cannot seriously think that every little chicken, that is hatched, has to rummage through all possible theories until it lights upon the good idea of picking up something and eating it. On the contrary, you think the chicken has an innate idea of doing this; that is to say, that it can think of this, but has no faculty of thinking anything else. The chicken you say pecks by instinct. But if you are going to think every poor chicken endowed with an innate tendency toward a positive truth, why should you think that to man alone this gift is denied? [41, 5.591].

The naturalistic view of instinct involves at least two aspects: *evolutionary/adaptive* and *perceptual*—as a 'certain insight' [41, 5.173]: the instinctual insight that leads to a hypothesis is considered by Peirce to be of 'the same general class of operations to which Perceptive Judgments belong'[16] (*ibid*.). Hence, Peirce considers the capacity to guess correct hypotheses as instinctive and enrooted

[15] Instinct is of course in part conscious: it is 'always partially controlled by the deliberate exercise of imagination and reflection' [41, 7.381].

[16] I have described perception as abduction in the subsection 1.2.

in our evolution and from this perspective abduction is surely a property of naturally evolving organisms:

> If you carefully consider with an unbiased mind all the circumstances of the early history of science and all the other facts bearing on the question […] I am quite sure that you must be brought to acknowledge that man's mind has a natural adaptation to imagining correct theories of some kind, and in particular to correct theories about forces, without some glimmer of which he could not form social ties and consequently could not reproduce his kind [41, 5.591].[17]

Peirce also says 'Thought is not necessarily connected with brain. It appears in the work of bees, of crystals, and throughout the purely physical world; and one can no more deny that it is really there, than that the colours, the shapes, etc., of objects are really there' [41, 4.551]. It is vital to explain the meaning of this important statement.

First of all it has to be noted that instincts themselves can undergo modifications through evolution: they are 'inherited habits, or in a more accurate language, inherited dispositions' [41, 2.170]. Elsewhere Peirce seems to maintain that instinct is not really relevant in scientific reasoning but that it is typical of just 'the reasoning of practical men about every day affairs'. So as to say, we can perform instinctive abduction (that is not controlled, not 'reasoned') in practical reasoning, but this is not typical of scientific thinking:

> These two [practical and scientific reasoning] would be shown to be governed by somewhat different principles, inasmuch as the practical reasoning is forced to reach some definite conclusion promptly, while science can wait a century or five centuries, if need be, before coming to any conclusion at all. Another cause which acts still more strongly to differentiate the methodeutic of theoretical and practical reasoning is that the latter can be regulated by instinct acting in its natural way, while theory of how one should reason depends upon one's ultimate purpose and is modified with every modification of ethics. Theory is thus at a special disadvantage here; but instinct within its proper domain is generally far keener, and surer, and above all swifter, than any deduction from theory can be. Besides, logical instinct has, at all events, to be employed in applying the theory. On the other hand, the ultimate purpose of pure science, as such, is perfectly definite and simple; the theory of purely scientific reasoning can be worked out with mathematical certainty; and the application of the theory does not require the logical instinct to be strained beyond its natural function. On the other hand, if we attempt to apply natural logical instinct to purely scientific questions of any difficulty, it not only becomes uncertain, but if it is heeded, the voice of instinct itself is that objective considerations should be the decisive ones.[18]

[17]Cognitive anthropologist Atran advocated a similar view about a century later, arguing in his *Cognitive Foundations of Natural History* that the evolution of religion and pre-scientific forms of knowledge into fully blown science could be accounted for just recurring to the concepts of *culture* and *cognition*, understanding the latter as 'the internal structure of ideas by which the world is conceptualized' [4, p. 3]. Peirce's philosophical speculations have been recently corroborated by a growing interest in *folk science*, i.e. in the study of uneducated expectations about natural aspects such as biology, mechanics, psychology, physiology and so on. Berlin and his colleagues pioneered the exploration of folkbiological expectations across different cultures [5]. The existence of folk science does not make the case for the actuality of a *lumen naturalis* predisposing humans towards Truth, but for the reality of a penchant (which is also at the level of perception) towards truthfulness: [23] argues that the success of science partially comes from 'the ways in which scientists learn to leverage understandings in other minds and to outsource explanatory work through sophisticated methods of deference and simplification of complex systems,' (p. 826) but such ways of relying on other people's knowledge in order to achieve better approximations of the truth about a matter are actually pre-existent in laypeople and children.

[18] Cf. Arisbe Website, http://members.door.net/arisbe/. The passage comes from MS L75 Logic, regarded as semeiotic (The Carnegie application of 1902).

I think that the considerations above do not mean, as some commentators seem to maintain [19, 40, 46], that instinct—as a kind of mysterious, not analysed, guessing power—'does not' operate at the level of conscious inferences as in the case of scientific reasoning. I think a better interpretation is the following that I am proposing here: certainly instinct, which I consider a simple and not a mysterious endowment of human beings, is at the basis of both 'practical' and scientific reasoning, in turn instinct shows the obvious origin of both in natural evolution. If every kind of cognitive activity is rooted in a hybrid interplay with external sources and representations, which exhibit their specific constraints and features, it does not appear surprising that '[…] the instincts conducive to assimilation of food, and the instincts conducive to reproduction, must have involved from the beginning certain tendencies to think truly about physics, on the one hand, and about psychics, on the other. It is somehow more than a *mere* figure of speech to say that nature fecundates the mind of man with ideas which, when those ideas grow up, will resemble their father, Nature' [41, 5.591]. Hence, from an evolutionary perspective instincts are rooted in humans in this interplay between internal and external aspects and so it is obvious to see that externalities ('Nature') 'fecundate' the mind. In this perspective abduction represents the most interesting fruit of this 'fecundated' mind.

Beyond the multifarious and sometimes contrasting Peircean intellectual strategies and steps in illustrating concepts like inference, abduction, perception and instinct, which of course are of great interest for the historians of philosophy,[19] the perspective I am describing here seems able to clearly focus on some central recent cognitive issues which I contend also implicitly underlie Peircean thoughts: nature fecundates the mind because it is through a disembodiment and extension of the mind in nature that in turn nature affects the mind. If we contend a conception of mind as 'extended', it is simple to grasp its instinctual part as shaped by evolution through the constraints found in nature itself. It is in this sense that the mind's abductive guesses—both instinctual and reasoned—can be classified as hypotheses 'akin to the truth' concerning nature and the external world because the mind grows up together with the representational delegations[20] to that 'nature' (external world) that the mind itself has made throughout the history of culture by constructing what some present-day biologists call cognitive niches.[21] In this strict perspective hypotheses are not merely made by pure *unnatural* chance.[22]

Peirce says, in the framework of his *synechism* that '[…] the reaction between mind and matter would be of not essential different kind from the action between parts of mind that are in continuous union' [41, 6.277]. This is clearly seen if we notice that '[…] habit is by no means a mental fact. Empirically, we find that some plants take habits. The stream of water that wears a bed for itself is forming a habit' [41, 5.492]. Finally, here the passage we already quoted above, clearly establishing Peirce's concerns about the mind: 'Thought is not necessarily connected with brain. It appears in the work of bees, of crystals, and throughout the purely physical world; and one can no more deny that it is really there, than that the colours, the shapes, etc., of objects are really there' [41, 4.551].

To conclude, seeing abduction as rooted in instinct vs. in inference represents a conflict we can overcome, following Peirce, simply by observing that the work of abduction is partly explicable as an instinctual biological phenomenon and partly as a 'logical' operation related to 'plastic' cognitive endowments of all organisms. I entirely agree with Peirce: a guess in science, the appearance of a

[19] For example, in the latest writings at the beginning of XX century Peirce more clearly stresses the instinctual nature of abduction and at the same time its inferential nature [40, p. 150]. On the various approaches regarding perception in Peircean texts cf. [56].

[20] Representational delegations are those cognitive acts that transform the natural environment in a cognitive one (a *cognitive niche*).

[21] Cf. [27, 28, 39]. I have illustrated in detail the concept of cognitive niche in ch. 6 of [32].

[22] This is not a view that conflicts with the idea of God's creation of human instinct: it is instead meant on this basis, that we can add, with Peirce, the theistic hypothesis, if desired.

new hypothesis, is also[23] a biological phenomenon and so it is related to instinct: in the sense that first of all we can analogize the appearance of a new hypothesis to a 'trustworthy' chance variation in biological evolution [41, 7.38], even if of course the evolution, e.g.—of scientific guesses does not conform to the pattern of biological evolution [9, p. 427]. An abduced hypothesis introduces a change (and a chance) in the semiotic processes to advance new perspectives in the co-evolution of the organism and the environment: it is in this way that they find a continuous mutual variation. The organism modifies its character in order to reach better fitness; however, the environment (already artificially—culturally—modified, i.e. a cognitive niche), is equally continuously changing and very sensitive to every modification. In summary, the fact that abduction is akin to truth is guaranteed at both the metaphysical and evolutionary levels: the case of instinct and the case of perception described by Peirce are striking, both provide abductions that are immediately and spontaneously generated but at the same time activated and efficacious, certainly not 'in sufferance' (as Woods would say, referring to the case of the standard activity of abducing hypotheses in natural science), and so in need of empirical evaluation.

2 The epistemology of evidentially inert knowledge enhancing

After having illustrated the philosophical and evolutionary justifications provided by Peirce to substantiate the truth-reliability of abductive cognition, some actual examples of knowledge enhancing abductions active in science, that nevertheless are evidentially inert,[24] have to be exhibited. To this aim the epistemological problem of guessing conventions and of guessing models in science will be illustrated in the following two subsections.

2.1 Guessing conventions in science. Abduction and strategic plausibility

We will now consider some aspects dealing with Poincaré's famous conventionalism of the principles of physics and the possibility of negating conventions. We will soon see that even if conventions in science are abduced and accepted, and their presence favours and increases knowledge, they are nevertheless not evidentially sensitive, so to say—at least in the sense that it is not possible to empirically falsify them. Consequently abduced conventions are evidentially inert but knowledge enhancing.

Gabbay and Woods maintain we can face a kind of abduction that, basically,

- is not plausibilist

at least in the sense of the word plausibilist indicated by the classical models provided in the literature. They say: 'It is not uncommon for philosophers to speak of the contribution made by the hypothesis of action-at-a-distance as one of explaining otherwise unexplainable observational data. […] Like numerous instances of D-N explanation, Newtonian explanations need convey no elucidation of their explicanda. They need confer no jot of further intelligibility to them. The action-at-a-distance equation serves Newton's theory in a wholly instrumental sense. It allows the gravitational theory to predict observations that it would not otherwise be able to predict' [15, pp. 118–119]. In my first book on abduction [30] I made some examples of abductive reasoning that basically are instrumentalist without clearly acknowledging it, later on I have adopted this concept in [32,

[23] Of course this conclusion does not mean that artefacts like computers do not or cannot perform abductions. The recent history of artificial intelligence in building systems able to perform diagnoses and creativity clearly illustrates this point.

[24] That is not inferences to the best explanation in the classical sense of the expression, involving an empirical evaluation phase.

ch. 2] where I noted that, e.g. in this case Newtonian explanations are seen as epistemically agnostic conjectures, i.e. they lack the classical epistemic virtues envisaged by the neopositivistic tradition. These abductions are secured by instrumental considerations and accepted because doing so enables one's target to be hit. They cannot be discharged because of their possible implausibility, and on the basis of empirical disconfirmation.

From the point of view of radical instrumentalist abduction this example is striking because it shows how these abduced principles fail all tests that would reveal them as having a traditional epistemic value, so that they are not subject to discharge except for their instrumental value. Again we are dealing with abductive cognitive fruits that are not strictly ignorance-preserving, and that not only mitigate ignorance but also provide further reliable knowledge.

An extension of Poincaré's so-called *geometric conventionalism*, according to which the choice of a geometry is only justified by considerations of simplicity, in a psychological and pragmatic sense ('commodisme'), is the *generalized conventionalism*, expressing the conventional character of the principles of physics:

> The principles of mathematical physics (for example, the principle of conservation of energy, Hamilton's principle in geometrical optics and in dynamics, etc.) systematize experimental results usually achieved on the basis of two (or more) rival theories, such as the emission and the undulation theory of light, or Fresnel's and Neumann's wave theories, or Fresnel's optics and Maxwell's electromagnetic theory, etc. They express the common empirical content as well as (at least part of) the mathematical structure of such rival theories and, therefore, can (but need not) be given alternative theoretical interpretations [17, pp. 27–28].

From the epistemological point of view it is important to stress that these abductively hypothesized conventional principles usually survive the demise of theories and are therefore responsible for the continuity of scientific progress: in a sense they show a radical instrumental character. Evidence from history of science, e.g. in the case of Mayer's establishment of the principle of conservation of energy [11], shows that conventional principles are usually fruit of complicated abductive processes and not of simple inductive generalizations. Moreover, they are not empirically falsifiable; as stated by Poincaré in *Science and Hypothesis*.[25]

Poincaré says 'we have a right to make' these conventions, and so their abduction is instantly knowledge enhancing.

> The principles of mechanics are therefore presented to us under two different aspects. On the one hand, they are truths founded on experiment, and verified approximately as far as almost isolated systems are concerned; on the other hand they are postulates applicable to the whole of the universe and regarded as rigorously true. If these postulates possess a generality and a certainty which the experimental truths from which they were deduced lack, it is because they reduce in final analysis to a simple convention that we have a right to make, because we are certain beforehand that no experiment can contradict it. This convention, however, is not absolutely arbitrary; it is not the child of our caprice. We admit it because certain experiments have shown us that it will be convenient, and thus is explained how experiment has built up the principles of mechanics, and why, moreover, it cannot reverse them [44, pp. 135–136].

Following Poincaré we can say that conventional principles of mechanics derive abductively from experience, as regards their 'genesis', but cannot be falsified by experience because they contribute

[25] Poincaré says that these principles are 'deduced' from experimental truths, but it is unlikely to think of them as fruit of deduction instead of abduction. I think that the use of the word deduction is just a way adopted by Poincaré to refer to a generic kind of scientific inference.

to 'constitute' the experience itself, in a proper Kantian sense. The experience has only suggested their adoption because they are *convenient*: there is a precise analogy with the well-known case of geometrical conventions, but also many differences, which pertain the 'objects' studied.

The conventional principles of mechanics should not be confused with geometrical conventions: 'The experiments which have led us to adopt as more convenient the fundamental conventions of mechanics refer to bodies which have nothing in common with those that are studied by geometry. They refer to the properties of solid bodies and to the propagation of light in a straight line. These are mechanical, optical experiments' [44, pp. 136–137], they are not, Poincaré immediately declares, '*des expériences de géométrie*' (*ibid.*): 'And even the probable reason why our geometry seems convenient to us is, that our bodies, our hands, and our limbs enjoy the properties of solid bodies. Our fundamental experiments are pre-eminently physiological experiments which refer, not to the space which is the object that geometry must study, but to our body—that is to say, to the instrument which we use for that study. On the other hand, the fundamental conventions of mechanics and experiments which prove to us that they are convenient, certainly refer to the same objects or to analogous objects. Conventional and general principles are the natural and direct generalizations of experimental and particular principles' (*ibid.*).

Poincaré continues:

> Principles are conventions and definitions in disguise. They are, however, derived from experimental laws, and these laws have, so to speak, been erected into principles to which our mind attributes an absolute value. Some philosophers have generalized far too much. They have thought that the principles were the whole of science, and therefore that the whole of science was conventional. This paradoxical doctrine, which is called nominalism, cannot stand examination. How can a law become a principle? [44, p. 138].

If the experimental laws of experimental physics are the source of the conventional principles themselves, conventionalism escapes nominalism.

As stated at the beginning of this subsection, conventional principles survive the demise (falsification) of theories in such a way that they underlie the incessant spectacle of scientific revolutions: 'It is the mathematical physics of our fathers which has familiarized us little by little with these various principles; which has habituated us to recognize them under the different vestments in which they disguise themselves' [45, p. 95]. Underlying revolutions of physics, conventional principles guarantee the historicity and the growth of science itself. Moreover, the conventional principles surely imply '[…] *firstly*, that there has been a *growing tendency* in modern physics to *formulate and solve* physical problems *within powerful, and more abstract, mathematical systems of assumptions* […]; *secondly*, the role of conventional principles has been growing and *our ability to discriminate experimentally between alternative abstract systems* which, with a great approximation, save the phenomena *has been diminishing* (by comparison to the testing of simple conjunctions of empirical generalizations)' [17, p. 28].

Moreover, as stated above, they are not empirically falsifiable: 'The principles of mechanics […] reduce in final analysis to a simple convention that we have a right to make, because we are certain beforehand that no experiment can contradict it' [44, p. 136].

Up to now I have considered in details how the conventional principles guarantee the revolutionary changes of physics and why they cannot be considered arbitrary, being motivated by—and abduced from—the *experimental laws* of the 'experimental physics', i.e. by experience. Even if arbitrary and conventional, the conventional principles too can be substituted by others. This is the main problem treated by Poincaré in the last passages of Chapter IX, 'The Future of Mathematical Physics', in *The Value of Science*. Already the simple case of 'linguistic' changes in science '[…] suffices to

reveal generalizations not before suspected' [45, p. 78]. By means of the new abductive discoveries, scientists arrive at a point where they are able to '[…] admire the delicate harmony of numbers and forms; they marvel when a new discovery opens to them an unexpected perspective' [45, p. 76], a new perspective that is always provisional, fallible, open to further confirmations or falsifications when compared to rival perspectives.

We have seen how the conventional principles of physics guarantee this continuous extension of experience thanks to the various perspectives and forms expressed by experimental physics. However, because conventional, 'no experiment can contradict them'. The experience only abductively suggested the principles, and they, since absolute, have become constitutive just of the empirical horizon common to rival experimental theories. These principles, even if abductively hypothesized just for strategic/instrumental reasons as important not *ad hoc* conventions, represent an interesting example of knowledge enhancing abduction evidentially inert: we have seen how the conventional principles of physics guarantee the continuous extension of experience thanks to the various perspectives and forms expressed by experimental physics and they cannot be dismissed on the basis of experimental tests.[26]

2.2 Knowledge enhancing through models. Abducing scientific models vs. abducing fictions

At the end of the subsection 1.3 above I have reminded that abduction is guessing and that human beings are very good at it. I also contended that it is in the case of empirical science that abduction seems to better exhibit its more representative epistemic virtues. In subsection 1.4, I have shown the classical Peircean philosophical and evolutionary justifications of that dominant character of abduction which is called 'ampliative' in the standard literature. Finally, in the previous subsection, I have illustrated the case of conventions, extremely important in physics, evidentially inert fruits of abduction—at least from the point of view of their impossible falsification—but nevertheless knowledge enhancing. I will examine in the present subsection that in science we do not have to consider the process of abducing models as a process of abducing fictions: as the reader can easily guess, this clarification will be intertwined with the aim of individuating other knowledge enhancing functions of abduction, even when clearly and immediately seen as evidentially inert.

Let us start with an example still provided by Woods, who illustrates the case of Planck's abduction as a case in which the epistemologist could see an active function of the so-called 'fictions':

> When in his quest for a unified treatment of the laws of black body radiation, Planck thought up the quantum hypothesis, it was a proposition for which there wasn't a shred of antecedent evidence and none at all abduced by its presence as antecedent in the subjunctive conditional on which its provisional conjecture was based. Planck thought that the very idea of the quantum was bereft of physical meaning. It is no condition on abductive adequacy that abduced hypotheses turn out well at experimental trial. There are more things whose truth was a reasonable thing to conjecture than actually turn out to be true. […] In some sense, the quantum hypothesis was down to Planck. Planck was the one who thought it up. Planck was the one who selected it for provisional engagement in a suitably adjusted physics. Some philosophers might see in these involvements a case for fictionalism [60].

'Planck was the one who thought it up', in this important creative event of the history of science.

[26] Beyond the aim of this study is the analysis of the problem of withdrawing the conventional principles of physics, which is solved by Poincaré by appealing to a form of weak negation. I have illustrated this issue in [32, ch. 2].

Not only in the case of key hypotheses like the one proposed by Planck, but also in the case of *models* that are built as 'epistemic mediators' inside a more extended process of scientific cognition, it is unlikely to admit they are abduced fictions, surely not in the minimal unequivocal sense of the word as it is adopted in the literary/narrative frameworks. Indeed, current epistemological analysis of the role models in science is often philosophically unproblematic and misleading. Scientific models are now not only considered as useful ways for explaining facts or discovering new entities, laws and theories, but are also rubricated under various new labels: from the classical ones, abstract entities and idealizations, to the more recent, fictions, credible worlds, missing systems, as make-believe, parables, as functional, as epistemic actions, as revealing capacities.[27] This proliferation of explanatory metaphors is amazing, if we consider the huge quantity of knowledge on scientific models that had already been produced both in epistemology and in cognitive science. Some of the authors involved in the debate on fictionalism are also especially engaged in a controversy about the legitimacy especially of speaking of fictions in the case of scientific models.

Even if evidentially inert in themselves, I think that the abduced models, both in scientific reasoning and in human perception (cf. subsection 1.4 above), cannot be considered as neither mere fictions, simple surrogates or make-believe, nor they are unproblematic idealizations. I am neither denying that models as idealizations and abstractions are a pervasive and permanent feature of science, nor that models, which are produced with the aim of finding the consequences of theories—often very smart and creative—are very important. I just stress that the 'fundamental' role played by models in science is the one we find in the core abductive discovery processes, and that these kinds of models cannot be indicated as fictional at all, because they are *constitutive* of new scientific frameworks and new empirical domains.[28] The abduction of these models in science is epistemically productive, models are just inert in the perspective of a *direct* empirical significance but they play a 'causal' role in generating it: scientific models can be empirically false, but they are not fictions, instead they are knowledge enhancing devices, which play an important role in reaching empirically fecund knowledge. It is clear here we are dealing with cases in which abduction is not ignorance-preserving.

Suárez [51] provides some case studies, especially from astrophysics and concerning quantum model of measurement, emphasizing the inferential function of the supposed to be 'fictional' assumptions in models: I deem this function to be ancillary in science, even if often highly innovative. Speaking of Thomson's plum pudding model, Suárez maintains that, basically 'The model served an essential pragmatic purpose in generating quick and expedient inference at the theoretical level, and then in turn from the theoretical to the experimental level. It articulated a space of reasons, a background of assumptions against which the participants in the debates could sustain their arguments for and against these three hypotheses' (p. 163). In these cases the fact that various assumptions of the models are empirically false is pretty clear and so is the 'improvement in the expediency of the inferences that can be drawn from the models to the observable quantities' (p. 165):[29] the problem is that in these cases models, however, are not fictions—at least in the minimal unequivocal sense of the word as it is adopted in the literary/narrative frameworks—but just the usual idealizations or

[27] An illustration of the main problems of fictionalism and a reference to the current literature on the subject is given in my recent [34].

[28] In this last sense the capacity of scientific models to constitute new empirical domains and so new *empirical knowability* is ideally related to the emphasis that epistemology, in the last century, put on the theory-ladenness of scientific facts (Hanson, Popper, Lakatos, Kuhn): in this light, the formulation of observation statements presupposes significant knowledge, and the search for new observability in science is guided by scientific modelling.

[29] It has to be added that Suárez does not certainly conflate scientific modelling with literary fictionalizing. He distinguishes scientific fictions from other kinds of fictions—the scientific ones are constrained by both the logic of inference and, in particular, the requirement to fit in with the empirical domain [51, 53]—in the framework of an envisaged compatibility of 'scientific fiction' with realism. This epistemological acknowledgement is not often present in other followers of fictionalism.

abstractions, already well-known and well-studied, as devices, stratagems, and strategies that lead to efficient results and that are not discarded just because they are not fake chances from the perspective of scientific rationality.[30] Two consequences derive:

- the role of models as 'expediency of the inferences' in peripheral aspects of scientific research, well-known from centuries in science, does not have to be confused with the *constitutive* role of modelling in the central *abductive* creative processes, when new conceptually more or less revolutionary perspectives are advanced;
- models are—so to say—just models that idealize and/or abstract, but these last two aspects have to be strictly criticized in the light of recent epistemologico/cognitive literature as special kinds of epistemic actions, as I have illustrated in more detail in [34, section 3.1]: abstractness and ideality cannot be solely related to empirical inadequacy and/or to theoretical incoherence [51, p. 168], in a static view of the scientific enterprise.

The considerations I have just illustrated show that model-based abduction in science is not truth preserving. Nevertheless, in the received view above, even if the guessed scientific models seem left in epistemic sufferance, scientific models cannot be considered works of fictions. At this point it can be said that one thing is to abduce fictions, like in the case of creations in literature, another thing is to abduce models in empirical science. Abducing fictions in literature is also certainly knowledge enhancing—like it is the case of scientific models—because we cannot surely imagine that literature does not provide knowledge of some kind. Moreover, how ignorance-preservation is at stake in these two cases? In the first case ignorance-preservation is related to an *esthetic* failure—a fictional character can be a literary failure, discarded by the author herself—in the second one to the possible experimental discredit, which would lead to the consequent lack of *rational* success of scientific enterprise. However, no need of using in the second case the word fiction: scientific models cannot be fictions.

However, we have also to remember that—normally—abductive processes to new concepts and models in literature and in science have also to be seen as a *continuum*—a sequence—of guessed hypotheses: in both cases if the production of an intermediate abductive hypothetical model fails (and so it is abandoned), this step *could still be seen as a significant cognitive achievement* if it is useful to provide some *crucial* new information later on, useful to produce further 'successful' hypotheses (e.g. to provide, respectively, Anna Karenina or Bohr' planetary model of the interior of atom). In this light we can say that the failed abductions were not completely esthetically or epistemically inert, because they facilitated the subsequent processes of hypothesis generation until the successful one.

2.2.1 Strategic abductive steps: dynamic vs. static view of scientific models

We have to further analyse in detail the process of abducing a sequence of possible hypotheses in a segment of scientific reasoning. This kind of process is often considered a *heuristic strategy*. For example, Hintikka stresses the strategic nature of abductive adoption of hypotheses inside a cognitive process:

> […] a strategic justification does not provide a warrant for any one particular step in the process. Such a particular step may not in any obvious way aid and abet the overall aim of the inquiry. For instance, such a step might provide neither any new information relevant to the aim of the inquiry nor any new confirmation for what has already been established, and yet might serve

[30] I discussed the role of chance-seeking in scientific discovery in [31].

> *crucially* the inquiry—for instance, by opening up the possibility of a question whose answer does so. Furthermore, notwithstanding the views of reliabilists, the idea of a non-strategic justification that they choose is not only mistaken but in the last analysis incoherent. From the theory of strategic processes misleadingly labelled game theory, it is known that what can be valuated (assigned 'utilities' to) are in principle only strategies, not particular moves. Hence a theory of epistemic processes that operates with 'warrants' for particular belief changes or other things that can be said of particular moves in our 'games' of inquiry is inevitably going to be unsatisfactory in the long run. One of the many things that Peirce's use of the term 'hypothesis' can serve to highlight is precisely the strategic character of any justification of abduction. Being strategic, such justification does not per se lend any reliability to the outcome of some particular abductive inference. This outcome has the status of a hypothesis. Whatever reliability it may possess has to be established by the inductive component of inquiry [18, pp. 57–58].

We can agree with Hintikka that it is certainly true that we cannot have 'warrants' at the level of strategic justification of particular steps in the process, and that it is also obviously true that the reliabilists are wrong suggesting the idea of a non-strategic justification. In my perspective it is the conclusion provided by Hintikka that is not satisfactory: 'Being strategic, such justification does not per se lend any reliability to the outcome of some particular abductive inference. This outcome has the status of a hypothesis. Whatever reliability it may possess has to be established by the inductive component of inquiry'. Hence, e.g. in the case of a scientific model abductively guessed, we would have to conclude—following Hintikka—that it is not reliable to the outcome of the cognitive process, indeed we have to wait for the empirical 'judgment'. Does this mean that every abductive guess is damned to be ignorance-preserving if evidentially inert? I do not think so.

To solve this problem a remark about the need of avoiding a confusion between a static and a dynamic view of scientific cognition has to be addressed. Indeed I think it is misleading to analyse the activity of abducting models in science by adopting a confounding and unclear mixture of static and dynamic aspects of the scientific enterprise. Temporal features of cognition count in understanding abduction. Scientific abduced models in a static perspective (e.g. when inserted in a textbook or in a text concerning history of empirical science) certainly appear—but just appear—justified 'by the inductive component of inquiry', i.e. only in the light of the successful empirical evaluation that has been finally performed: in this case the *instrumental*—and fruitful—character of the abduced models becomes manifest, but their *constitutive* function disappears.

Please imagine that you are isolating a single moment of a dynamic creative scientific process, forgetting what will happen or happened later on, in the last case, e.g. as it is testified in a historical narrative or in a textbook. Contrarily to the previous static view, some—the creative ones—of the abduced scientific models, once seen inside the living dynamics of scientific cognition[31] actually appear to be *explicit*, *reproducible* and *constitutive* machineries built and manipulated to the gnoseological aim of reaching a final overall scientific result empirically evaluated, a result *not yet available*: but we have a result, the only result we have is the intermediate one, the model. The final knowledge just results not yet available because the target system and its complicated experimental apparatuses have *not yet* been built. The problem is that our snapshot of the single moment shows to us that these final outcomes will be built *only* thanks to the gift in terms of subsequent *knowability* provided by the intermediate models themselves. In few words: the models at play are *creative*, because they positively 'establish' the root that leads to the empirical success. In a sense, their creativity 'is' their reliability, they do not need further strategic or not strategic

[31] Which, by the way, is the key topic of epistemology at least since Karl Popper and Thomas Kuhn.

justification or warrants. If we do not acknowledge this—Kantian, I would say—aspect, we are not able to befittingly and honestly understand what abduction is as a knowledge enhancing cognitive device.

When Hintikka contends that the abductive steps which lead to intermediate models *cannot* have 'warrants' at the level of strategic justification, and also at the level of non-strategic justification, in my perspective we can relieve ourselves of this burden of epistemic sufferance just acknowledging we are dealing with *creative* models. If we only see models in empirical science in the light of the future achieved empirical success we obviously see them just as *provisional* guesses, *devoid* of justification and still and intrinsically looking for it. On the contrary, they are occasionally justified by themselves—abductively—just because creative, and so *constitutive* of a fruitful epistemic 'cognitive travel'. In sum, coming back to the main issue we are dealing with in this article, those models are sometimes knowledge enhancing at their level, even if *locally* evidentially inert.

Let us reconsider in this perspective the problem of models as fictions. If we consider that the abduced models—in science—are fictions they are certainly evidentially inert; unfortunately they would be also ignorance-preserving, because they will lack, as fictions, the capacity to produce a kind of intermediate knowledge endowed with the *epistemological virtue* of rationality. Woods furnishes a further crucial insight on the superfluity of speaking of fictions in science adopting a useful distinction between what he calls infinite (forlon) falsehoods and fictions. Given the fact that fictions denotate and infinite falsehoods do not, infinite falsehoods cannot be fictions. At the same time his strict argumentation provides a further justification of the knowledge enhancing status of scientific models. He thinks that the detonation question for infinite falsehoods (e.g. models that involve infinite populations in biology, that, as I have already said, many epistemologists consider fictions)

> […] is a trick question. It is a logical commonplace that, unlike truth, falsity is not preserved under consequence. How surprising can it be, then, that when $T \vdash O_i$ holds, the falsity embedded in T is not passed on to the O_i? The very fact of T's empirical adequacy precludes the detonation of its falsities. It is precisely here that fictionality's explosiveness achieves a grip. Since detonation is not a problem for falsely tinctured T_s, fictions are not required to fix it. Yet if fictions were called into play, they would create a denotation problem for T, and would guarantee that it could not be solved. For, again, detonation precludes empirical adequacy. I take this to be a serious discouragement of the fictionalist programme for science [60].

Again, if the O_i of an empirically adequate T are not derivable

> […] in the absence of $T's$ infinitely remote falsehoods, then $T's$ connection to those O_i cannot be grounding. T cannot be said to have demonstrated those consequences or to have provided a reason that supports them. This is a puzzle. But suppose, now, that fictions were called into play with a view to solving it. Then T wouldn't be empirically adequate. (Fictionality detonates.) The grounding question asks how T can be empirically adequate if it doesn't lend grounding support to the O_i in virtue of which this is so. But if fictions are let loose here, the empirical adequacy of T is lost. The grounding question wouldn't arise (*ibid.*)

Another—more epistemologically oriented—explanation of the issue at stake is given by Kuorikoski and Lehtinen [26, p. 121], who contend that: 'The epistemic problem in modelling arises from the fact that models always include false assumptions, and because of this, even though the derivation within the model is usually deductively valid, we do not know whether our model-based inferences reliably lead to true conclusions'. The problem is that false premises (also due to the presence in models of both substantive and auxiliary assumptions) are not exploited in the

cognitive process, because, in various heuristic steps, only the *co-exact* ones are exploited. For example, the notion of co-exact proprieties, introduced by Manders [35], is worth to be further studied in fields that go beyond the realm of discovery processes of classical geometry, in which it has been nicely underscored. Mumma [38, p. 264] illustrates that Euclid's diagrams contribute to proofs only through their co-exact properties. Indeed 'Euclid never infers an exact property from a diagram unless it follows directly from a co-exact property. Exact relations between magnitudes which are not exhibited as a containment are either assumed from the outset or are proved via a chain of inferences in the text. It is not difficult to hypothesize why Euclid would have restricted himself in such a way. Any proof, diagrammatic or otherwise, ought to be reproducible'.

Moreover, as I have already noted, some false—eventually abduced—assumptions are considered as such *only if* seen in the light of the still 'to be known' target system, and so they appear false only in a *post hoc* analysis, but they are perfectly true—and so knowledge enhancing—in the model itself in its relative autonomy during the smart heuristic cognitive process related to its exploitation. Falsities in one perspective can be considered (local) truths in another one. So various aspects of the model are the legitimately true basis for the subsequent exploration of its behaviour and performance of further abductions or other inferences to statements concerning the target system. I agree with Morrison: 'I see this not as a logical problem of deriving true conclusions from false premises but rather an epistemic one that deals with the way false representations transmit information about concrete cases' [37, p. 111].[32]

Morrison [37] is certainly not inclined to see models as fictions because she emphasizes that in science they are specifically related to ('finer graded') ways of understanding and explaining 'real systems', far beyond their more collateral predictive capabilities and their virtues in approximating. She indeed further clarifies that the models which is appropriate to label as *abstract* resist—in the so-called process of de-idealization—corrections or relaxing of the unrealistic assumptions (such as in the case of mathematical abstractions or when models furnish the sudden chance for the applicability of equations), because they are 'necessary' to arrive to certain results: models are not redundantly required for the derivation of the predictions by which gain a contact with the observational level. The fact that in these models 'relevant features' are subtracted to focus on a single—and so isolated—set of properties or laws, as stressed by Cartwright [8], is not their central quality, because what is at stake is their capacity to furnish an overall new depiction of an empirical (and/or theoretical, like in case of mathematics or logic) framework: '[…] We have a description of a physically unrealizable situation that is required to explain a physically realizable one' (p. 130). In a similar vein Woods nicely concludes 'A central role for empirically forlorn representations in model-based science is the establishment of non-probative premise–conclusion linkages in ways that set up their conclusions for empirical negotiation at the checkout counter' [59]. The cognitive situations just described still reflect those cases of abductive successes in model-based reasoning, which are simply not probatively successful from the ultimate empirical point of view, but, as we are trying to demonstrate, not necessarily devoid of a knowledge enhancing status.

Moreover, many models, easier to define, which is better to classify as *idealizations*, allow '[…] for the addition of correction factors that bring the model system closer (in representational terms)

[32]Further information about the problem of the mapping between models and target systems through *interpretation* are provided by Contessa [10, p. 65]—interpretation is seen as more fundamental than surrogative-reasoning: 'The model can be used as a generator of hypotheses about the system, hypotheses whose truth or falsity needs to be empirically investigated'. By using the concept of interpretation (analytically and not hermeneutically defined) the author in my opinion also quickly adumbrates the creative aspects in science, that coincide with the fundamental problem of model-based and manipulative abduction.

to the physical system being modelled or described' [37, p. 111]. It is, e.g. the case of simple pendulum, where we know how to add corrections to deal with concrete phenomena. Idealizations distort or omit properties, instead abstractions introduce a specific kind of representation 'that is not amenable to correction and is necessary for explation/prediction of the target system' (p. 112), and which provides information and transfer of knowledge. Morrison's characterization of scientific models as abstract is in tune with my emphasis on models as creative and so accepted as *constitutive*, beyond the mere role played by models as idealizations, which instead allow corrections and refinements.[33] In this perspective, 'abstract' models, either related to prepare and favour mathematization or directly involving mathematical tools, have to be intended as poietic ways of producing new intelligibilty of the essential features of the target systems phenomena, and not mere expedients for facilitating calculations. If idealization *resembles* the phenomena to be better understood, abstract models abductively *constitute* the resemblance itself.

It is not that 'fictions provide inferential shortcuts in models; and the fact that this is the main or only reason for their use distinguishes them as fictional' [53, p. 239], even if Vaihinger would agree with this functionalist perspective on fictions.[34] Indeed, even if it is not decisive to say 'that the inferential characterization provides a way to distinguish precisely scientific from non-scientific uses of fiction', the abduced models used in non-scientific practices may also trigger inferences, and the problem here is more fundamental. In science, models are not used and intended as fictions, they are just labelled as fictions because of a juxtaposition of some recent philosophers of science, who certainly in this way render the scientific enterprise more similar to other more common modes of human cognition: after all fictions are ubiquitous in human cognition, and science is a cognitive activity like others. Unfortunately science never aimed to abduce 'fictions' at the basic levels of its activities, so that the recent fictionalism does not add new and fresh knowledge about the status of models in science, and tends to obfuscate the distinctions between different areas of human cognition, such as science, religion, arts and philosophy. In the end, 'epistemic fictionalism' tends to enforce a kind 'epistemic concealment', which can obliterate the actual gnoseological finalities of science, shading in a kind of debate about entities and their classification that could remind of medieval scholasticism. Abduction is a widespread cognitive activity, like its logical models teach to us, and so it is the activity of abducing models, but this huge extension paradoxically furnishes a further reason for the philosopher, the epistemologist and the cognitive scientist to study, differentiate and respect the various types of knowledge, beliefs and levels of truth and/or rationality more or less involved.

2.2.2 Mathematics and manipulative abduction

In this subsection I plan to illustrate some arguments devoted to further substantiate my analysis of the limitations of the ignorance-preserving view of abduction. Abduction patently enhances knowledge also in mathematics, where it is completely manifest that every successful abductive process is evidentially inert. Still in this case my *eco-cognitive model* (EC-model) of abduction is useful (cf. above subsection 1.2), because it aims at providing a concluding explanation of the cognitive machineries that can render abduction knowledge enhancing even in absence of empirical evaluation.

Mathematical discoveries are not empirically negotiable and also in the case of mathematics 'models' are at play. Elsewhere [32, ch. 3] I called the external scientific models 'mimetic' (e.g. a

[33] On the constitutive vs. descriptive role of models cf. also [50].

[34] I have to add that Suárez does not defend the view according to which models are fictions: even if he defends the view that models contain or lead to fictional assumptions, he rejects the identification of models and fictions, preferring instead to stay 'quietist' about the ontology of models, and focusing rather on modelling as an activity—see in particular [52].

geometrical diagram drawn on a blackboard or an in-vitro model of neurobiological process),[35] not in a military sense, as camouflaged tools to trick the hostile eco-human systems, but just as structures that mimic the target systems for epistemic aims. In this perspective I described the centrality of the so-called 'disembodiment of the mind' in the case of semiotic cognitive processes occurring in science. Disembodiment of the mind refers to the cognitive interplay between internal and external representations, *mimetic* and, possibly, *creative*, where the problem of the continuous interaction between on-line and off-line (e.g. in inner rehearsal) intelligence can properly be addressed.[36]

I consider this interplay critical in analysing the relation between meaningful semiotic internal resources and devices and their dynamical interactions with the externalized semiotic materiality already stored in the environment (scientific artefactual models, in our case). This external materiality plays a specific role in the interplay due to the fact that it exhibits (and operates through) its own cognitive constraints. Hence, minds are 'extended' and artificial in themselves. It is at the level of that continuous interaction between on-line and off-line intelligence that I underlined the importance of what I called *manipulative abduction*.

In this eco-cognitive perspective on abduction—also recalled in the first part of this article—it is worth to note that, among the advantages offered by the external models in their role of perturbing the internal ones, there are not only the unexpected features that can be offered thanks to their intrinsic materiality—wonderfully embedded in models of empirical science—but also more neutral yet fruitful devices. These more neutral devices can be, e.g. exemplified thanks to the case of the re-representations in internal mental models of already externalized mathematical symbols, which at the first sight we can label, again, 'fictions': 'Apparently the brain immediately translates a positive integer into a mental representation of its quantity. In contrast, symbols that represent non-intuitive concepts remain partially semantically inaccessible to us, we do not reconstruct them, but use them as they stand' [12]. For example, it is well-known that Leibniz adopted the notation *dx* for the infinitesimals he genially introduced, and called them *fictions bien fondées*, given their semantic paradoxical character: they lacked a referent in Leibnizian infinitesimal calculus, but were at the basis of plenty of new astonishing mathematical results.[37] De Cruz and De Smedt call this property of symbols 'semantic opacity', which renders them underdetermined, allowing further creative abductive processes where those same symbols can be relatively freely exploited in novel contexts for multiple cognitive aims. Semantic opacity favours a kind of reasoning that is unbiased by those intuitive aspects that possibly involve stereotypes or intended uncontrolled interpretations, typical of other less opaque external models/representations.

Peirce too was clearly aware, speaking of the model-based aspects of deductive reasoning, that there is an 'experimenting upon this image [the external model/diagram] in the imagination', where the idea that human imagination is always favoured by a kind of prosthesis, the external model as an 'external imagination', is pretty clear, even in case of classical geometrical deduction: '[…]

[35]On the related problem of resemblance (similarity, isomorphism, homomorphism, etc.) in scientific modelling see [34, section 3.5].

[36]In this eco-cognitive perspective we can even more strongly agree with Morrison's when she is pretty clear about the excessive habit of labelling fictional scientific models simply because they are superficially seen as 'unrealistic': 'Although there is a temptation to categorize any type of unrealistic representation as a 'fiction', I have argued that this would be a mistake, primarily because this way of categorizing the use of unrealistic representations tells us very little about the role those representations play in producing knowledge' [37, p. 133].

[37]To confront critiques and suspects about the legitimacy of the new number *dx*, Leibniz prudently conceded that *dx* can be considered a fiction, but a 'well founded' one. The birth of non-standard analysis, an 'alternative calculus' invented by Abraham Robinson [48], based on infinitesimal numbers in the spirit of Leibniz's method, revealed that infinitesimals are not at all fictions, through an extension of the real numbers system $\mathbb{R}$ to the system $\mathbb{R}^*$ containing infinitesimals smaller in the absolute value than any positive real number. On this story cf. the recent [22].

namely, deduction consists in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of experimenting upon this image in the imagination and of observing the result so as to discover unnoticed and hidden relations among the parts' [41, 3.363].

Also Hintikka clearly shows the 'embarrassing' presence of fruitful abductive creative moments in deduction, which are invaded by strategic hypothetical interventions crucial to proceed and reach the final results, and that of course are evidentially inert. Also in deduction, the presence of abductive events coincides with their knowledge enhancing character: here too these strategic aspects reflect the pure—productive—conjectural element of abductive inference and its capacity to guessing right. Hintikka clearly points out the abductive nature of the inferential phase in which the existential quantifier is introduced. This case is in turn related to his emphasis on the *strategically* positing of the 'right questions' which 'depend on one's ability to anticipate their answers' [18, p. 55]:

> […] the very same sentence can serve as the presupposition of a question and as the premise of a deductive step. For instance, an existential sentence of the form
> (1) $(\exists x)S[x]$
> can serve either as the presupposition of the question
> (2) What (who, when, where,...), say x, is such that $S[x]$?
> or as the premise of an existential instantiation that introduces a John Doe—like 'dummy name' of an 'arbitrary name', say β. In the former case, the output of the relevant step is a sentence of the form
> (3) $S[b]$
> where b is a singular term—for instance, a proper name. In the latter case, the output is of the form
> (4) $S[\beta]$
> Here, (4) differs from (3) only by having a dummy name, whereas in (3) there was a real name. […] It seems to me that Peirce had an intuitive understanding of this type of similarity between abductive and deductive inferences. […] These similarities between questions (abductive steps) and logical inferences (deductive steps) are purely formal, however. An epistemological assimilation of the two to each other on the mere basis of such formal similarities would be irresponsible. The crucial insight here is that behind these formal similarities there lies a remarkable strategic similarity [18, pp. 53–54].

The strategic similarity resorts to the need of the reasoner of using one of the propositions that are available to her as presuppositions or as premises, *both* in the case of abduction and of deduction: 'Which sentence or sentences should I use as the premise or as the premises of a deductive inference? It can be shown that the most sensitive strategic question in deduction is: Which sentence should I use first as the premise of an existential instantiation or its generalization, functional instantiation? […] If the inquirer is reasoning empirically (interrogatively), the next strategic question is: Which one of the available sentences should I use as the presupposition of a why question? These candidate sentences are the very same ones that could be used as premises of existential instantiations, suitably generalized. Neither question admits in general of a mechanical answer, in the sense that there is in neither case any recursive function that always specifies an optimal choice. […] In this sense, the strategic principles of abductive reasoning, interpreted as I have done, *are the same as the strategic principles governing deduction*' ([18, p. 54], emphases added).

Taking advantage of the previous notes about the abductive aspect of mathematical reasoning (and discovery) and logical deduction let us come back to the problem of abduction and of its presumptive ignorance-preserving character. I have many times stressed in my works that manipulative abduction

(I have just quoted few lines above), which is widespread in scientific reasoning, is a process in which a hypothesis is formed resorting to a basically extra-theoretical and extra-sentential behaviour that aims at creating communicable accounts of new experiences to the final aim of integrating the successful results into previously existing systems of experimental and linguistic (theoretical) practices. Manipulative abduction represents a kind of redistribution of the epistemic and cognitive effort to manage objects and information that cannot be immediately represented or found internally. An example of manipulative abduction is exactly the case of the human use of the construction of external models, e.g. in a neural engineering laboratory, useful to make observations and 'experiments' to transform one cognitive state into another to discover new properties of the target systems. Manipulative abduction refers to those more unplanned and unconscious action-based cognitive processes I have characterized as forms of 'thinking through doing'.[38] It is clear that manipulative abduction in science basically deals with the handling of external models in their intertwining with the internal ones. Consequently, even if related to experiments occasionally performed with the help of external models sometimes mediated by artefacts, manipulative abduction has to be considered—obviously in mathematics but also in the case of empirical science, evidentially inert, even if of course not necessarily ignorance-preserving, as I have tried to demonstrate in this article.

I have contended that manipulative abduction is also active in mathematics. For example, we have already seen that Peirce, in the case of mathematics, speaking of the model-based aspects of this kind of deductive reasoning, hypothesized there is an 'experimenting upon this image [the external model/diagram] in the imagination', so showing how human geometrical imagination is always triggered by a kind of prosthesis, the external model as an 'external imagination'. Analogously, taking advantage of a fictional view on models and of the pretence theory Frigg [14, p. 266 ff.] interestingly sees imagination as an authorized intersubjective game of make-believe sanctioned by the 'prop' (an object, e.g. material models, movies, paintings, plays, etc.) and its rules of generation. This theory also works as a metaphor of abductive processes, in terms of some concepts taken from the theory of literary and artistic fictions: again, I think that it is neither necessary to adopt a fictionalist view in the case of science, nor the pretence theory adds something relevant to the issue and, moreover, this kind fictionalism would obscure the knowledge enhancing role of abduction we are describing in this article.

Analogously, in the example I have illustrated in [34], concerning the exploitation of concrete/external models (e.g. in-vitro or computational) in a scientific lab, scientists do not pretend anything and are not engaged in the relative make-believe process, if not in the trivial sense that almost every human intersubjective interplay can be seen as such. The in-vitro networks of cultured neurons of that case or the Peircean Euclidean diagram used by the ancient Greek geometers are just the opposite of a mere fiction or of a generic make-believe interplay, they are instead more or less mimetic (possibly creative and so enhancers of new knowledge not already available) external models—reached through manipulative abduction—which are expected to provide reliable information about the target system. They aim at abductively discovering some new representations about the neurons in the first case and about the pure concepts of geometry in the second.

The reason of my skepticism about the vision of models in terms of the pretence theory can be illustrated taking advantage of some classical but still astounding theses derived from classical Kantian philosophy and Thom's mathematical semiophysics. Immanuel Kant was clearly aware of the interplay between internal and external models—exemplified in the case of a formal science like

[38] I have to note that manipulative abduction also happens when we are *thinking through doing* (and not only, in a pragmatic sense, about doing). This kind of action-based cognition can hardly be intended as completely intentional and conscious.

mathematics—as an example of genuine knowledge production (and, occasionally, of discovery). In his transcendental terms, Kant says that in geometrical construction '[…] I must not restrict my attention to what I am actually thinking in my concept of a triangle (this is nothing more than the mere definition); I must pass beyond it to properties which are not contained in this concept, but yet belong to it' [21, A718-B746, p. 580]. Hence, for Kant models in science (in this case, of geometry) are first of all *constructions* that go beyond what the researcher simply 'thinks'. We have seen that manipulative abduction is a kind of, usually model based, abduction that exploits external models endowed with delegated (and often implicit) cognitive roles and attributes: (1) The model is external and the strategy that organizes the manipulations is unknown a priori. (2) The result achieved is new (if we, for instance in this geometrical case, refer to the constructions of the first creators of geometry), and adds properties not contained before in the concept (the Kantian to 'pass beyond' or 'advance beyond' the given concept [21, A154-B194, p. 192]).[39]

Iconicity is central for Peirce, who analogously to Kant, maintains that '[…] philosophical reasoning is reasoning with words; whereas theorematic reasoning, or mathematical reasoning, is reasoning with specially constructed schemata' [41, 4.233]; moreover, he uses diagrammatic and schematic as synonyms, thus relating his considerations to the Kantian tradition where schemata mediate between intellect and phenomena.[40] The following is the famous-related passage in the *Critique of Pure Reason* ('Transcendental Doctrine of Method'):

> Suppose a philosopher be given the concept of a triangle and he be left to find out, in his own way, what relation the sum of its angles bears to a right angle. He has nothing but the concept of a figure enclosed by three straight lines, and possessing three angles. However long he meditates on this concept, he will never produce anything new. He can analyse and clarify the concept of a straight line or of an angle or of the number three, but he can never arrive at any properties not already contained in these concepts. Now let the geometrician take up these questions. He at once begins by constructing a triangle. Since he knows that the sum of two right angles is exactly equal to the sum of all the adjacent angles which can be constructed from a single point on a straight line, he prolongs one side of his triangle and obtains two adjacent angles, which together are equal to two right angles. He then divides the external angle by drawing a line parallel to the opposite side of the triangle, and observes that he has thus obtained an external adjacent angle which is equal to an internal angle—and so on. In this fashion, through a chain of inferences guided throughout by intuition, he arrives at a fully evident and universally valid solution of the problem [21, A716-B744, pp. 578-579].

Here 'intuition' is the Kantian word that expresses our present reference to what we call 'external model'.

We can depict the situation of the philosopher described by Kant at the beginning of the previous passage also taking advantage of some ideas coming from the catastrophe theory. As a human being who is not able to produce anything new relating to the angles of the triangle, the philosopher experiences a feeling of frustration (just like Kölher's monkey which cannot keep the banana out of reach). The bad affective experience 'deforms' the organism's regulatory structure by complicating it and the cognitive process stops altogether. The geometer instead 'at once constructs the triangle'

[39] Of course in the case we are using diagrams to demonstrate already known theorems (for instance in didactic settings), the strategy of manipulations is often already available and the result is not new.

[40] Schematism, a fruit of the imagination is, according to Kant, '[…] an art concealed in the depths of the human soul, whose real modes of activity nature is hardly likely ever to allow us to discover, and to have open to our gaze' [21, A141-B181, p. 183]. Now we have at our disposal, thanks to epistemology and cognitive science, a lot of knowledge about the cognitive processes which correspond to Kantian schematism. On models as epistemic mediators in mathematics cf. [7].

[the scientist constructs the model] i.e. he makes an external representation of a triangle and acts on it with suitable manipulations. Thom thinks that this action is triggered by a 'sleeping phase' generated by possible previous frustrations which then change the cognitive status of the geometer's available and correct internal idea of triangle (like the Kantian philosopher, he 'has nothing but the concept of a figure enclosed by three straight lines, and possessing three angles', but his action is triggered by a sleeping phase). Here the idea of the triangle is no longer the occasion for 'meditation', 'analysis' and 'clarification' of the 'concepts' at play, like in the case of the 'philosopher'. Here the inner concept of triangle—symbolized as insufficient—is amplified and transformed thanks to the sleeping phase (a kind of Kantian imagination active through schematization) in a prosthetic triangle to be put outside, in some external support. The instrument (here an external diagram) becomes the extension of an organ, a cognitive *anchorage*:

> What is strictly speaking the end […] [in our case, to find the sum of the internal angles of a triangle] must be set aside in order to concentrate on the means of getting there. Thus the problem arises, a sort of vague notion altogether suggested by the state of privation. […] As a science, heuristics does not exist. There is only one possible explanation: the affective trauma of privation leads to a folding of the regulation figure. But if it is to be stabilized, there must be some *exterior form* to hold on to. So this *anchorage problem* remains whole and the above considerations provide no answer as to why the folding is stabilized in certain animals or certain human beings whilst in others (the majority of cases, needless to say!) it fails [55, pp. 63–64] (emphasis added).[41]

As I have anticipated, here we see that, even in the mathematical discovery processes, manipulative abduction is based on the interplay between internal and external representations (not only diagrams, but also written proofs, etc.): the final result is an abductive hypothesis which assumes the clothes of a Kantian 'stipulation', endowed with epistemic virtues, that same 'productive' stipulation, squarely evidentially inert, we have seen at work in the case of conventions and models in empirical science.

Hence, in my eco-cognitive perspective models in science cannot be considered neither abstract (in the traditional ambiguous sense) nor fictional and abducing them is not necessarily ignorance-preserving: after all, I have already pointed out that scientists do not have any intention to propose fictions, instead they provide models as tools that reshape a generic cognitive niche as an epistemological niche to the aim of performing a genuine struggle for 'rationally' representing the external world. Models, the war machines used in this struggle, which I have called *epistemic warfare* [34], to stress the determined—strictly epistemic—dynamism of the adopted tools that are at stake, are not illusional fictions or stratagems used for example to cheat nature or swindle human beings, but just concrete, unambiguous and well-disposed tactical abductive intermediate weapons—endowed with epistemic virtues—able to strategically 'attack' nature (the target systems) to further unveil its structure. The kind of cognition provided by models in scientific enterprise is meant to increase 'rational knowledge' and has to be distinguished by the kind of cognition that is, e.g. at play in moral judgments or in the case of rhetorical stratagems at the level of what I call 'military intelligence', which characterizes the struggle between (and inside) human collectives of various kind.[42] In turn, fictions in works of fictions are, e.g. meant to unveil human life and characters in new esthetic

[41] A full analysis of the Kölher's chimpanzee getting hold of a stick to knock a banana hanging out of reach in terms of the mathematical models of the perception and the capture catastrophes is given in [55, pp. 62–64]. On the role of emotions, e.g. frustration, in scientific discovery cf. [54].

[42] The 'military' function of abduction in the framework of the so-called *coalition enforcement* is analysed in [33].

perspectives and/or to criticize them through a moral teaching, whereas fictions and stratagems in wars are meant to trick the enemy and possibly destroy the eco-human targets.

I contend that epistemologists and researchers in abductive cognition must not forget that various cognitive processes present a 'military' nature, even if it is not evident in various aspects and uses of syntactilized human natural language and in abstract knowledge. It is hard to directly see this 'military intelligence' in the many *epistemic* functions of natural language, e.g. when it is simply employed to transmit scientific results in an academic laboratory situation, or when we gather information from the Internet—expressed in linguistic terms and numbers—about the weather. However, we cannot forget that even the more abstract character of knowledge packages embedded in certain uses of language (and in hybrid languages, like in the case of mathematics, which involves considerable symbolic parts) still plays a significant role in changing the moral behaviour of human collectives. For example, the production and the transmission of new scientific knowledge in human social groups not only operates on information but also implements and distributes roles, capacities, constraints and possibilities of actions. This process is intrinsically moral because in turn it generates precise distinctions, powers, duties and chances which can create new between-groups and in-group violent (often) conflicts, or reshape older pre-existent ones.[43]

It is important to explicitly emphasize the intrinsic relativity of the status of concepts like truth, rationality, knowledge, ignorance: their reciprocal entanglement tends to reciprocally depict the respective meanings. Here an example: it has to be said that successful abductions that are performed at the moral level I have mentioned above immediately acquire a deontological status. They are *epistemically inert* even if they increase something that we can certainly call 'knowledge': the 'moral' knowledge human individuals need in a given situation. The use of the word knowledge depicts the meaning of the word ignorance: at least under the perspective of the last moral case, the abductions involved *are not* ignorance-preserving, because do not preserve the subjective moral ignorance in front of the problems of 'military intelligence' at stake. Nevertheless, the abduced hypothetical knowledge in the case of these—primarily moral—endeavours can be easily seen a piece of 'false' knowledge, from the empirical and/or rational point of view, but still active and efficient, and in this case we are legitimated to call the involved abductions as basically ignorance-preserving. Indeed, common moral knowledge of beings like us is not intrinsically truth-sensitive.[44]

Final, coming back to the problem of model-based abduction, I confess I cannot see how we can speak of the creative activity of abducing in the case of the ideal pendulum in the same way we speak of abducing the character of Anna Karenina or the visual mental model we can adopt to initiate a moral deliberation, where often false cognitive contents are at play but still efficient: it seems to me that, by adopting a confusion of the different levels at play, we are running the risk of inadvertently opening the gates of epistemology to a kind of relativistic post-modernism *à la mode*, even if fictionalists seem to avoid this possible confusion by producing—often useful—taxonomies about the slight differences between fictions in science and in other cognitive practices. I am aware of the fact that 'epistemological' fictionalism does not consider fictions forgery or fake, i.e. something 'far from being execrable', instead, something 'we cherish' [14, p. 249], but to say that scientific and literary fictions are both 'good' fictions is a bit of a theoretical oversimplification, because it is science that created—thanks to the fact abduction is akin to truth—beyond literature and poetry,

[43] I am deriving the expression 'military intelligence' from René Thom [55], who relates it to the role played by language and cognition in the so-called *coalition enforcement*, i.e. at the level of their complementary effects in the affirmation of moralities and related conducts, and the consequent perpetration of possible violent punishments.

[44] I have illustrated the role of abduction in military intelligence in [33, ch. 2], where I have extendedly treated the relationship between cognition and violence.

new kinds of models committed to a specific production of rationality, constitutively aiming at not being fictional, but instead truth-making.[45]

Galileo is explicitly clear about the distinction between science (he calls 'philosophy' in the following celebrated passage) and literary fiction:

> In Sarsi[46] I seem to discern the firm belief that in philosophizing one must support oneself upon the opinion of some celebrated author, as if our minds ought to remain completely sterile and barren unless wedded to the reasoning of some other person. Possibly he thinks that philosophy is a book of fiction by some writer, like the Iliad or Orlando Furioso, productions in which the least important thing is whether what is written there is true. Well, Sarsi, that is not how matters stand. Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth [16, pp. 237–238].[47]

3 Conclusion

The status of abduction is very controversial. When dealing with abductive reasoning misinterpretations and equivocations are common. What did Peirce mean when he considered abduction both a kind of inference and a kind of instinct or when he considered perception a kind of abduction? Does abduction involve only the generation of hypotheses or their evaluation too? Are the criteria for the best explanation in abductive reasoning epistemic, or pragmatic, or both? Does abduction preserve ignorance or extend truth or both? The article has tried to answer these questions centring the attention to the so-called ignorance-preservation character of abduction, such as it is illustrated by the GW-model of abduction.

I have contended that certainly abductive reasoning is a response to an ignorance-problem. Nevertheless, through abduction, knowledge can be enhanced, even when abduction is not considered an IBE in the classical sense of the expression, i.e. an inference necessarily characterized by an empirical evaluation phase. To study this conundrum I exploited my EC-model of abduction and illustrated three cases taken from the areas of both philosophy and epistemology, in which abduction de facto more or less creatively produces knowledge enhancing effects even when, I repeat, it is not intended as an IBE which involves the 'inductive evaluation', as Peirce called it. I have first of all described Peirce's *evolutionary* and *metaphysical* ideas, which aim at confirming that abduction is constitutively *akin to truth*, even if certainly always ignorance-preserving or mitigating, at least in the sense that the 'absolute truth' is never reached. I have also addressed the case of abducing *conventions* in empirical science, which highlights a clear example of abductive generation of

[45] Cf. the example of the famous thought experiment of Galileo concerning the problem of falling bodies, I have described in [34, section 3.6].

[46] Lothario Sarsi of Siguenza is the pseudonym of the Jesuit Orazio Grassi, author of *The Astronomical and Philosophical Balance*. In *The Assayer*, Galileo weighs the astronomical views of Orazio Grassi about the nature of the comets, and finds them wanting [16, p. 231].

[47] As Bertolotti [6] observes, the quotation obviously should not be used as an authority weapon against those who advocate the fictional nature of scientific models, because we would commit a fallacy, given the fact that to affirm that scientific models are fictions does not coincide with saying that the whole scientific endeavour has a fictional nature. Thus, the use of this quotation does not aim at getting definitively rid of fictionalism through the authority of one of the founding fathers of modern science.

new scientific knowledge even if the reached hypotheses remain evidentially inert, together with an analysis of the abduction in generating scientific models, which play fundamental pre-empirical 'rational' knowledge enhancing roles. Finally, I have also shown that scientific models in a static perspective can appear fictional to the epistemologists, but their fictional character vaporizes if a dynamic perspective is assumed. Abduction in scientific model-based reasoning is not a questionable process of guessing fictions.

Acknowledgements

For the instructive criticisms and precedent discussions and correspondence that helped me to develop my critique of the ignorance-preserving character of abduction and my analysis of fictionalism, I am indebted and grateful to John Woods, Mauricio Suárez, Shahid Rahman, Alirio Rosales and Tommaso Bertolotti.

References

[1] A. Aliseda. *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. PhD Thesis, Institute for Logic, Language and Computation, Amsterdam, 1997.
[2] A. Aliseda. *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Springer, 2006.
[3] N. Angius. Towards model-based abductive reasoning in automated software testing. *Logic Journal of the IGPL*, 2013 (this issue).
[4] S. Atran. *Cognitive Foundations of Natural History: Towards an Anthropology of Science*. Cambridge University Press, 1990.
[5] B. Berlin, D. Breedlove and P. Raven. General principles of classification and nomenclature in folk biology. *American Anthropologist*, **74**, 214–242, 1973.
[6] T. Bertolotti. From mindless modelling to scientific models. The case of emerging models. In *Philosophy and Cognitive Science. Western and Eastern Studies*, L. Magnani and P. Li, eds, pp. 75–104. Springer, 2012.
[7] M. J. Boumans. Mathematics as quasi-matter to build models as instruments. *Probabilities, Laws, and Structures*, In M. Weber, D. Dieks, W. J. Gonzalez, S. Hartman, F. Stadler and M. Stöltzner, eds, pp. 307–318. Springer, 2012.
[8] N. Cartwright. *Nature's Capacities and Their Measurement*. Oxford University Press, 1989.
[9] V. Colapietro. Conjectures concerning an uncertain faculty. *Semiotica*, **153**, 413–430, 2005.
[10] G. Contessa. Scientific representation, interpretation, and surrogative reasoning. *Philosophy of Science*, **74**, 48–68, 2007.
[11] J. Coopersmith. *Energy, the Subtle Concept: The Discovery of Feynman's Blocks from Leibniz to Einstein*. Oxford University Press, 2010.
[12] H. De Cruz and J. De Smedt. Mathematical symbols as epistemic actions. *Synthese*, 2011, doi:10.1007/s11229-010-9837-9.
[13] P. Feyerabend. *Against Method*. Verso, 1975.
[14] R. Frigg. Models and fiction. *Synthese*, **172**, 251–268, 2010.
[15] D. M. Gabbay and J. Woods. *The Reach of Abduction*. North-Holland, 2005. Vol. 2 of *A Practical Logic of Cognitive Systems*.
[16] G. Galilei. *The Assayer* [1623]. In *Discoveries and Opinions of Galileo*, S. Drake, translated and ed., pp. 231–280. Doubleday, 1957.
[17] J. Giedymin. *Science and Convention. Essays on Henri Poincaré's Philosophy of Science and the Conventionalist Tradition*. Pergamon Press, 1982.

[18] J. Hintikka. *Socratic Epistemology. Explorations of Knowledge-Seeking by Questioning*. Cambridge University Press, 2007.
[19] M. H. G. Hoffmann. Problems with Peirce's concept of abduction. *Foundations of Science*, **4**, 271–305, 1999.
[20] A. Kakas, R. A. Kowalski and F. Toni. Abductive logic programming. *Journal of Logic and Computation*, **2**, 719–770, 1993.
[21] I. Kant. *Critique of Pure Reason*. MacMillan, 1929. Translated by N. Kemp Smith, originally published 1787, reprint 1998.
[22] M. G. Katz and D. Sherry. Leibniz's infinitesimals: their fictionality, their modern implementations, and their foes from Berkeley to Russell and beyond. *Erkenntnis*, 2012, doi:10.1007/s10670-012-9370-y.
[23] F. Keil. The feasibility of folk science. *Cognitive Science*, **34**, 826–862, 2010.
[24] R. A. Kowalski. *Logic for Problem Solving*. Elsevier, 1979.
[25] T. A. F. Kuipers. Abduction aiming at empirical progress of even truth approximation leading to a challenge for computational modelling. *Foundations of Science*, **4**, 307–323, 1999.
[26] J. Kuorikoski and A. Lehtinen. Incredible worlds, credible results. *Erkenntnis*, **70**, 119–131, 2009.
[27] K. N. Laland, F. J. Odling-Smee and M. W. Feldman. Niche construction, biological evolution and cultural change. *Behavioral and Brain Sciences*, **23**, 131–175, 2000.
[28] K. N. Laland, F. J. Odling-Smee and M. W. Feldman. Cultural niche construction and human evolution. *Journal of Evolutionary Biology*, **14**, 22–33, 2001.
[29] A. Mackonis. Inference to the best explanation, coherence and other explanatory virtues. *Synthese*, **190**, 975–995, 2013.
[30] L. Magnani. *Abduction, Reason, and Science. Processes of Discovery and Explanation*. Kluwer Academic/Plenum Publishers, 2001.
[31] L. Magnani. Abduction and chance discovery in science. *International Journal of Knowledge-Based and Intelligent Engineering*, **11**, 273–279, 2007.
[32] L. Magnani. *Abductive Cognition. The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*. Springer, 2009.
[33] L. Magnani. *Understanding Violence. The Interwining of Morality, Religion, and Violence: A Philosophical Stance*. Springer, 2011.
[34] L. Magnani. Scientific models are not fictions. Model-based science as epistemic warfare. In *Philosophy and Cognitive Science. Western and Eastern Studies*, L. Magnani and P. Li, eds, pp. 1–38. Springer, 2012.
[35] K. Manders. The Euclidean diagram. In *Philosophy of Mathematical Practice*, P. Mancosu, ed., pp. 112–183. Clarendon Press, 2008.
[36] J. Meheus, L. Verhoeven, M. Van Dyck and D. Provijn. Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In *Logical and Computational Aspects of Model-Based Reasoning*, L. Magnani, N. J. Nersessian and C. Pizzi, eds, pp. 39–71. Kluwer Academic Publishers, 2002.
[37] M. Morrison. Fictions, representations, and reality. In *Fictions in Science: Philosophical Essays on Modeling and Idealization*, M. Suárez, ed., pp. 110–135. Routledge, 2009.
[38] J. Mumma. Proofs, pictures, and Euclid. *Synthese*, **175**, 255–287, 2010.
[39] F. J. Odling-Smee, K. N. Laland and M. W. Feldman. *Niche Construction. The Neglected Process in Evolution*. Princeton University Press, 2003.
[40] S. Paavola. Peircean abduction: instinct or inference? *Semiotica*, **153**, 131–154, 2005.

[41] C. S. Peirce. *Collected Papers of Charles Sanders Peirce*. Harvard University Press, 1931–1958; vol. 1–6, C. Hartshorne, and P. Weiss, eds; vol. 7–8, A. W. Burks, ed.
[42] C. S. Peirce. Visual cognition and cognitive modelling. In *Philosophical Writings of Peirce*, J. Buchler, ed., pp. 302–305. Dover, 1955.
[43] C. S. Peirce. *Historical Perspectives on Peirce's Logic of Science: A History of Science*, C. Eisele.ed., vol. I–II. Mouton, 1987.
[44] H. Poincaré. *La science et l'hypothèse*. Flammarion, 1902. English translation by W. J. G. [only initials indicated], 1958, *Science and Hypothesis*, with a Preface by J. Larmor, The Walter Scott Publishing Co., 1905. Also reprinted in *Essential Writings of Henri Poincaré*, Random House, 2001.
[45] H. Poincaré. *La valeur de la science*. Flammarion, 1905. English translation by G. B. Halsted, 1958, *The Value of Science*, Dover Publications. Also reprinted in *Essential Writings of Henri Poincaré*, Random House, 2001.
[46] N. Rescher. Peirce on abduction, plausibility, and efficiency of scientific inquiry. In *Essays in the History of Philosophy*, N. Rescher, ed., pp. 309–326. Avebury, 1995.
[47] F. D. Rivera and J. Rossi Becker. Abduction-induction (generalization) processes of elementary majors on figural patterns in algebra. *Journal of Mathematical Behavior*, **26**, 140–155, 2007.
[48] A. Robinson. *Non-Standard Analysis*. North Holland, 1966.
[49] M. Shanahan. Perception as abduction: turning sensory data into meaningful representation. *Cognitive Science*, **29**, 103–134, 2005.
[50] A. Stefanov. Theoretical models as representations. *Journal for General Philosophy of Science*, **43**, 67–76, 2012.
[51] M. Suárez. Scientific fictions as rules of inference. In *Fictions in Science: Philosophical Essays on Modeling and Idealization*, M. Suárez, ed., pp. 158–178. Routledge, 2009.
[52] M. Suárez, (ed). *Fictions in Science: Philosophical Essays on Modeling and Idealization*. Routledge, 2009.
[53] M. Suárez. Fictions, inference, and realism. In *Fictions and Models: New Essays*, J. Woods, ed., pp. 225–245. Philosophia Verlag, 2010.
[54] P. Thagard. The passionate scientist: emotion in scientific cognition. In *The Cognitive Basis of Science*, P. Carruthers, S. Stich and M. Siegal, eds, pp. 235–250. Cambridge University Press, 2002.
[55] R. Thom. *Esquisse d'une sémiophysique*. InterEditions, 1988. Translated by V. Meyer, *Semio Physics: a Sketch*. Addison Wesley, 1990.
[56] C. Tiercelin. Abduction and the semiotic of perception. *Semiotica*, **153**, 389–412, 2005.
[57] J. Woods. Ignorance, inference and proof: abductive logic meets the criminal law. In *The Rules of Inference: Inferentialism in Law and Philosophy*, G. Tuzet and D. Canale, eds, pp. 151–185. Egea, 2009.
[58] J. Woods. Recent developments in abductive logic. *Studies in History and Philosophy of Science*, **42**, 240–244, 2011. Essay Review of L. Magnani, *Abductive Cognition. The Epistemologic and Eco-Cognitive Dimensions of Hypothetical Reasoning*, Springer, 2009.
[59] J. Woods. Mathematizing knowledge, 2012. Unpublished Paper.
[60] J. Woods. Against fictionalism. In *Model-Based Reasoning in Science and Technology. Theoretical and Cognitive Issues*, L. Magnani, ed., Springer, 2013 (forthcoming).
[61] J. Woods. *Errors of Reasoning. Naturalizing the Logic of Inference*. College Publications, London, 2013 (forthcoming).

Received 7 January 2013