Studies in Applied Philosophy,
Epistemology and Rational Ethics

SAPERE

John R. Shook
Sami Paavola *Editors*

Abduction in Cognition and Action

Logical Reasoning, Scientific Inquiry, and Social Practice

Springer

Studies in Applied Philosophy, Epistemology and Rational Ethics

Volume 59

Studies in Applied Philosophy, Epistemology and Rational Ethics (SAPERE) publishes new developments and advances in all the fields of philosophy, epistemology, and ethics, bringing them together with a cluster of scientific disciplines and technological outcomes: ranging from computer science to life sciences, from economics, law, and education to engineering, logic, and mathematics, from medicine to physics, human sciences, and politics. The series aims at covering all the challenging philosophical and ethical themes of contemporary society, making them appropriately applicable to contemporary theoretical and practical problems, impasses, controversies, and conflicts. Our scientific and technological era has offered "new" topics to all areas of philosophy and ethics – for instance concerning scientific rationality, creativity, human and artificial intelligence, social and folk epistemology, ordinary reasoning, cognitive niches and cultural evolution, ecological crisis, ecologically situated rationality, consciousness, freedom and responsibility, human identity and uniqueness, cooperation, altruism, intersubjectivity and empathy, spirituality, violence. The impact of such topics has been mainly undermined by contemporary cultural settings, whereas they should increase the demand of interdisciplinary applied knowledge and fresh and original understanding. In turn, traditional philosophical and ethical themes have been profoundly affected and transformed as well: they should be further examined as embedded and applied within their scientific and technological environments so to update their received and often old-fashioned disciplinary treatment and appeal. Applying philosophy individuates therefore a new research commitment for the 21st century, focused on the main problems of recent methodological, logical, epistemological, and cognitive aspects of modeling activities employed both in intellectual and scientific discovery, and in technological innovation, including the computational tools intertwined with such practices, to understand them in a wide and integrated perspective. **Studies in Applied Philosophy, Epistemology and Rational Ethics** means to demonstrate the contemporary practical relevance of this novel philosophical approach and thus to provide a home for monographs, lecture notes, selected contributions from specialized conferences and workshops as well as selected Ph.D. theses. The series welcomes contributions from philosophers as well as from scientists, engineers, and intellectuals interested in showing how applying philosophy can increase knowledge about our current world. Initial proposals can be sent to the Editor-in-Chief, Prof. Lorenzo Magnani, lmagnani@unipv.it:

- A short synopsis of the work or the introduction chapter
- The proposed Table of Contents
- The CV of the lead author(s).

For more information, please contact the Editor-in-Chief at lmagnani@unipv.it.

Indexed by SCOPUS, zbMATH, SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/10087

John R. Shook · Sami Paavola
Editors

Abduction in Cognition and Action

Logical Reasoning, Scientific Inquiry, and Social Practice

Springer

Editors
John R. Shook
Philosophy and Science Education
University at Buffalo
Amherst, NY, USA

Sami Paavola
Faculty of Educational Sciences
University of Helsinki
Helsinki, Finland

ISSN 2192-6255 ISSN 2192-6263 (electronic)
Studies in Applied Philosophy, Epistemology and Rational Ethics
ISBN 978-3-030-61772-1 ISBN 978-3-030-61773-8 (eBook)
https://doi.org/10.1007/978-3-030-61773-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Introduction

This book gathers together in one volume fresh essays offered from the frontiers of research into the logic and practice of abduction. The publication of this collected volume could not be more timely. In many ways, abduction has become established and essential to many fields beyond logic and philosophy of science. However, at the same time, the provocative implications of abduction and the precise workings of abductive thinking offer more and more puzzles and problems.

Abductive inference, with its emphasis on anticipatory guessing, predictive learning, and imaginative hypothesis, broadens the scope of the cognitivist repertoire for understanding human reasoning and action. As cognitivism views intelligence in general, knowledge's growth must involve cognitive processes for learning. Rejecting cognitivism entirely is not a fruitful agenda, since inference, fast or slow, is involved with any information processing and animal learning more complex than mere associative conditioning and pattern recognition.

If cognitivism in a stricter sense assumes that cognition's work is a separate matter apart from anything else the organism is interacting with, abduction presents a serious challenge. As pragmatism has reiterated, little about intelligence is comprehensible after presuming that thought occurs independently from other life activities. When abduction is taken seriously as a vital mode of inference, it is far easier to understand the full power of cognition, as a functional stage within a larger process of learning with various kinds of engagements to the environment.

Abduction in its primary sense is a mode of inference. Inference is truly worthy of careful and exhaustive study. Taking inference seriously, it is found to be pervasive across learning: sensible intuitions arise from fast or hidden inferences; memorization would be chaotic without inferential checks for consistency; and cogent rationalizing requires inferential links between facts and convictions. For humans, explicit inferential reasoning that modifies belief and leads toward knowledge is quite different from intuiting, memorizing, or rationalizing. Intuition presents factual matters for belief, but their plausibility relies on intimate familiarity. Memorization accepts unfamiliar facts on trust for expanding belief. Rationalization seeks relevant facts to support established belief. Inference for learning transcends all three by seeking out facts beyond the familiar and favored that may strengthen the credibility of a hesitant guess. Where inference goes, abduction may become involved.

In human reasoning, good reasoning should be conducive to knowledge. Proper deduction and induction have been upheld as high standards of justifying certain and probable knowledge. Since abduction is the source of original hypotheses, why should abduction be omitted from the analysis of knowledge? Intelligence at its most strenuous tries to acquire new knowledge, placing high demands on inference that mere intuition, memorization, and rationalization can never satisfy. To never go beyond available facts is to never learn something new. Inferential learning, unlike simpler modes, cannot take for granted the paired sides of inference: What shall be the relevant facts, and what may be a reasonable idea? Both parts will change during novel learning.

Strict cognitivism, by supposing that deduction and induction exhaust the main modes of explicit inference, reaches its dead-end here. Neither deduction nor induction in their classic form can yield adequate accounts of novel learning. Valid deduction may enlighten the deducer, but valid deduction in themselves is not enlarging knowledge itself. The probabilities of induction go a little farther than accumulated facts, but correlations cannot by themselves engender conceptions of unobserved causes. A third primary mode of hypothetical inference must be at work with original knowledge about anything beyond what may be evident and already accepted. This third mode was labeled as abduction by scientist and philosopher Charles S. Peirce originally in the 1860s (although then with a different name), and it played a major role in most everything philosophical that he produced during the rest of his life.

Abduction as a distinctive and legitimate mode of inference was overlooked by both empiricisms and rationalisms throughout the eighteenth and nineteenth centuries and ignored as just diversionary or fallacious by research on scientific methods during much of the twentieth century. That neglect is no longer the case. A familiarity with abduction is now common across many disciplines, and denials of abduction's significance for methodology, knowledge advancement, and scientific reasoning are heard far less frequently. Beyond logic itself, many areas of philosophy are reinvigorated by the study of abduction and its contributions to cognition, such as epistemology, practical reasoning, creativity, action theory, and decision theory. Abduction also has a central place in semiotics, linguistics, communications, rhetoric, and philosophy of language.

The cognitive sciences have accordingly been influenced heavily, especially where cognitive psychology and its research into perceptual recognition and belief revision are intersecting with computational sciences devoted to information theory, computational logic, artificial intelligence, machine learning, distributed intelligence systems, and robotics. Behavioral and social sciences are also attending to abductive reasoning where educational methods, agent decision theory, research on practices, and interpreting methodologies are key. Empirical research concerning a variety of types of abductive processes has been slow to emerge, but now quickly growing due to interest from so many fields.

In parallel with this growing attention to abduction across various disciplines, there has also been a proliferation of puzzles about abduction. Differing interpretations and applications for abduction have arisen, which is a welcome sign of its conceptual fertility. Yet many fundamental questions about abduction remain open. How is abduction specifically manifested in human cognition and intelligence? Are there different types of abduction, and if so, how might they be classified and related? Would abduction be expected to proceed quite differently in application within in different areas of research? What are the concrete roles and responsibilities attached to abduction during the course of empirical inquiry and scientific discovery? Can abduction be a logic of discovery? Can it be a valid mode of inference? How abduction is manifested in practical processes of inquiry? How does abduction actually work for various sorts of methodologies that relate to the conduct of human practices?

These sorts of key issues and questioning opportunities are taken up in this volume's chapters. Starting with Peirce, how should his original formulations of abduction receive accurate elaborations? Peirce's voluminous amount of texts has left room for various interpretatations and applications of abduction. Novel conceptions of abduction suggest revised understandings of other views on thought and learning in turn. Going further, alternative conceptions of abduction have been formulated, extending and surpassing Peirce's writings on abduction. Some chapters present new methods and tools for formalizing abduction, of potential use in schematizations for thinking, problem-solving, learning, and skill improvement. Presently, there are a variety of interpretations of abduction which can be connected with ways of classifying distinctive forms of abduction. These interrelationships soon prove to become quite complex, going beyond abduction in itself, to explore mixed forms of reasoning and inferences involved with a broad range of practical activities.

In past decades, much of the logical and philosophical literature about abduction concentrated on comparing schematizations of abductive inference with formalizations for deductive and inductive inferences and evaluating criteria of validity or degrees of probable truth. These are still important issues, and some chapters take notice of them. Beyond schematics, most of the chapters concentrate more on the dynamics of abduction applied for inventive enterprises through various means. This dynamic outlook on abduction permits further questions to be raised about abduction's roles in imagination, discovery, and creativity.

This volume covers the "state-of-the-art" opportunities in this interdisciplinary area of research on abduction. Its fourteen chapters are composed by a total of nineteen authors or co-authors, working in eleven countries around the world. Our contributors have long been at the forefront of discussions on abduction, and they now offer their updated approaches to the issues that they regard as central to abduction's contemporary relevance. All important areas of research about abduction are addressed, from its origins with C. S. Peirce, its broad relevance to inferential reasoning and scientific methods, and its enrichment of core problems in the cognitive and behavioral sciences. Parts of the book cover interlinked interests, but they are grouped under four bigger themes.

Part One on "Analyzing Abduction" presents papers about different sorts of abduction where they are most frequently found. In Chapter "Are there Types of Abduction? An Inquiry into a Comprehensive Classification of Types of Abduction," Jorge Alejandro Flórez explores ways of classifying types of abduction. He discusses Peirce's hesitance of making subdivisions of abduction and contrasts this to later suggestions about various types of abduction. He concludes by adding his own classification about interesting kinds of abduction based on abductive syllogisms. In Chapter "Practical Abduction for Research on Human Practices: Enriching Rather Than Testing a Hypothesis," Sami Paavola interprets abduction in relation to studies on human practices. He argues that the Peircean idea of abduction, as the first phase of inquiry that calls for further testing, is not the most apt idea for practice studies. Practical experimentation follows a different kind of logic closer to the Deweyan idea of a working hypothesis, or alternatively the method of ascending from the abstract to the concrete where abduction has a continuous role throughout the research process. In Chapter "Abductive Spaces: Modeling Concept Framework Revision with Category Theory," Rocco Gangle, Gianluca Caterina, and Fernando Tohmé describe how their model of abduction is interpretable as a representation of a community trying to retain commitments through modifications that appear to be needed for dealing with obstacles. This model offers a way to understand reasonable processes available for creative problem-solving by a social group.

Part Two on "Developing and Applying Peirce's Abduction" focuses on interpreting Peirce's abduction and relating it to other scholars' work. In Chapter "Abduction as "Leading Away": Aristotle, Peirce, and the Importance of Eco-cognitive Openness and Situatedness," Lorenzo Magnani discusses the background of situated abduction that can be discerned in Aristotle's texts. Magnani's own eco-cognitive model (EC-Model) of abduction enlarges on that theme with his "optimization of eco-cognitive openness and situatedness." More attention should be paid to "discoverability" and "diagnosticability" in abductive results. In Chapter "Methodeutic of Abduction," Francesco Bellucci and Ahti-Veikko Pietarinen analyze the meaning of "methodeutic" for understanding Peircean abduction and Peirce's conception of pragmatism. They argue that methodeutic for Peirce contributes essentially to the security of abduction, and they present three maxims as the bedrock of Peirce's methodeutic of abduction: experientiality for security, simplicity for uberty, and economy for advantageousness of reasoning. In Chapter "Peirce, Russell and Abductive Regression," John Woods contrasts Peirce and Bertrand Russell on the search for logical foundations of mathematics. He explores the relevance of abduction for philosophy of mathematics, in the context of an early twentieth-century debate over logicism. A resemblance between Peirce on abduction and Russell's "regressive method" to find premises of mathematics may not be as close as it seems.

Part Three on "Abduction, Mathematics, and Logic in Creative Discovery" analyzes and applies abduction with regard to creativity and creative reasoning. In Chapter "The Place of Logic in Creative Reason," Atocha Aliseda argues on behalf of the thesis that logic and creativity do not exclude each other. This concerns especially the possibility and practice of ampliative, or synthetic reasoning—that is, abduction. Aliseda sorts through Peirce's primary conceptions of abduction and his later views on the possibility of the logic of discovery. In Chapter "On Abducing the Axioms of Mathematics," Woosuk Park sharply distinguishes abduction from inference to the best explanation (IBE), in preparation for a discussion of mathematical discovery. Park disputes the idea that IBE explains the role of the axiom of choice in mathematics. By contrast, abduction supplies a better account of discovering and justifying mathematical axioms. In Chapter "Abduction, Complex Inferences, and Emergent Heuristics of Scientific Inquiry," John Shook discriminates complex modes of abduction needed for creative scientific discovery. As Peirce indicated, scientific methods of postulation and justification combine iterations of deduction, induction, and abduction, which together heighten the credibility of well-tested hypotheses. This "procedural abduction" attains levels of heuristics, higher than pseudo-science, which justify scientific realism about postulated entities. In Chapter "Abduction, The Logic of Scientific Creativity, and Scientific Realism," John Shook turns to the role of scientific communities which creatively re-conceive hypothesized entities during stages of inquiry that apply procedural abduction. According to the methodology of procedural abduction, any dichotomy between the logic of discovery and the logic of justification is dissolved. In consequence, the abductive methodology of scientific communities and the credible realism of scientific entities are mutually dependent.

Part Four on "Human Reasoning and Theories of Cognition" concentrates on novel interpretations of abduction in relation to theories of human cognition and reasoning. In Chapter "Abduction and the Logic of Inquiry: Modern Epistemic Vectors," Jay Schulkin reminds us how thinking is undertaken by a brain optimizing for predictive accuracy and successful activity. Problem-solving and adaptive behaviors guided by the plasticity of neural functioning would naturally take the form of abductions, performed under constraints of time and energy availability. In Chapter "Abductive Inference, Self-Knowledge, and the Myth of Introspection," Eric Charles and Nicholas Thompson analyze Peircean critique of the Cartesian approach to introspection. They argue that our self-knowledge requires similar inferential processes in order to attain judgments about the minds of others. These processes are, in an important sense, results from abductive inference. In Chapter "The Role of Collections of Objects in Abduction," Patricia Turrisi offers her deliberations about the meaning of objects of thought, especially material objects for abductive investigation. She gives various examples of collecting activity and resulting collections where (material) objects are central. These are examples of Peircean *logica utens*, that is, acts of reasoning without formal logical systematization. In Chapter "Our Themes on Abduction in Human Reasoning: A Synopsis," Emmanuelle-Anna Dietz Saldanha, Steffen Hölldobler, and Luís Moniz Pereira take for their topic the evident gap between actual human reasoning and classical logic. Rather than abandoning logic as a way to model how people think, cognitive theory should be supplemented by abduction. Their original Weak Completion Semantics makes better sense of the ways that people manage to make reasonable sense.

John R. Shook
jshook@pragmatism.org

Sami Paavola
sami.paavola@helsinki.fi

Contents

Contributors

Atocha Aliseda Instituto de Investigaciones Filosóficas, National Autonomous University of Mexico, México, Mexico

Francesco Bellucci University of Bologna, Bologna, Italy

Gianluca Caterina Endicott College, Beverly, MA, USA

Eric Phillip Charles American University, Washington, USA

Emmanuelle-Anna Dietz Saldanha Faculty of Computer Science, Dresden, Germany

Jorge Alejandro Flórez Restrepo Universidad de Caldas, Manizales, Colombia

Rocco Gangle Endicott College, Beverly, MA, USA

Steffen Hölldobler Faculty of Computer Science, Dresden, Germany; North-Caucasus Federal University, Stavropol, Russian Federation

Lorenzo Magnani Department of Humanities, Philosophy Section and Computational Philosophy Laboratory, University of Pavia, Pavia, Italy

Sami Paavola Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

Woosuk Park Korea Advanced Institute of Science and Technology, Daejeon, South Korea

Luís Moniz Pereira NOVA Laboratory for Computer Science and Informatics, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

Ahti-Veikko Pietarinen Tallinn University of Technology, Tallinn, Estonia; Research University Higher School of Economics, Moscow, Russia

Jay Schulkin Department of Neuroscience, Georgetown University, Washington, USA

John R. Shook University at Buffalo, Buffalo, NY, USA

Nicholas S. Thompson Clark University, Worcester, USA

Fernando Tohmé UNS-CONICET, Bahia Blanca, Argentina

Patricia Turrisi University of North Carolina, Wilmington, USA

John Woods University of British Columbia, Vancouver, BC, Canada

Analyzing Abduction

Are There Types of Abduction? An Inquiry into a Comprehensive Classification of Types of Abduction

Jorge Alejandro Flórez Restrepo

Abstract This chapter enquires into the possibility of having types of abduction. In the two first sections, it explores the classification of types of abduction given by Peirce (although he paradoxically argued as well that abduction has no divisions) and those given by scholars, such as Thagard, Eco, Hoffman, Schurz, Aliseda, and others. The journey across their suggestions of types of abduction permits not only to see the number and definition of those types, but it also permits to identify the criteria according to which those classifications were obtained. In the last section, the author evaluates those suggestions and unifies the criteria used by the scholars in order to propose that abduction is a very prolific and varied inference.

It is paradoxical that Peirce claimed both that there are two types of abduction and that abduction has no divisions. Most of the time, he asserted that there was just one type of abduction. In his early writings he analyzed logical inferences from a syllogistic point of view and he found that abduction only has a unique form in which a Result and Rule are premises and they conclude a Case; and in his late writings he never included types of abduction because he used what I call the categorical principle of degeneracy to achieve those classifications. Abduction was always considered by him as a firstness and therefore, it is not subject of further classifications; it remains as single and unique. Instead, deduction considered as a secondness can be divided into two types (theorematic and collorarial), and induction considered as a thirdness can be divided into three types (rudimentary, qualitative, and quantitative). Even his further divisions make deduction have three types, induction six types, but abduction remains always with a unique form.

J. A. Flórez Restrepo (✉)
Universidad de Caldas, Manizales, Colombia
e-mail: jorgealejandro.florez@ucaldas.edu.co

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_1

This is contradictory with his assertions in two different passages that abduction has two types depending upon conditions of time and consciousness. It is also contradictory with the proposal of several scholars who from the 1980s to nowadays have proposed a multitude of types of abductions.[1]

This article, then, is an inquiry that attempts to resolve this paradox. Its aim firstly is to evaluate and to expose Peirce's arguments to postulate mainly that abduction is an inference that has no types, and also to expose those two passages in Peirce's late writing (1903 and 1909) where he explored the possibilities of having two types of abduction. The second part of this text is intended to evaluate the arguments of contemporary scholars in order to postulate types of abductions. It is not my intention to defend Peirce tenaciously for even he is open to reevaluating his position due to his fallibilism. Instead, the inquiry that this article offers on the topic of types of abductions is open to arguments for and against. Therefore, in the third section, I will explore a theory of types of abduction based on the information acquired in the two first sections.

1 Peirce on Types of Abduction

Charles S. Peirce considered that "the chief business of the logician is to classify arguments; for all testing clearly depends on classification" (EP1, 186).[2] In fact, one of the branches of logic that he developed the most was Speculative Grammar, which observes the nature of signs and classifies their types. He attempted several classifications of signs and inferences during his life, but a constant characteristic of those classifications is that abduction has no subdivisions. Instead, induction and deduction have different subdivisions and in all those attempts of classification, he hesitated among different names and numbers for those subdivisions.

The classificatory aim of logic was exposed since his first article The New List of Categories (1868). There Peirce proposed the following classification of signs (included inferences):

---

[1]For a complete list of different proposals of types of abduction, see Paavola 2006.

[2]The references to Peirce's work are abbreviated as follows:

CP *Collected Papers of Charles Sanders Peirce*, edited by C. Hartshorne, P. Weiss (volumes 1–6), and A. Burks (volumes 7–8) (Cambridge, Mass.: Harvard University Press, 1931-58), followed by volume and paragraph number.

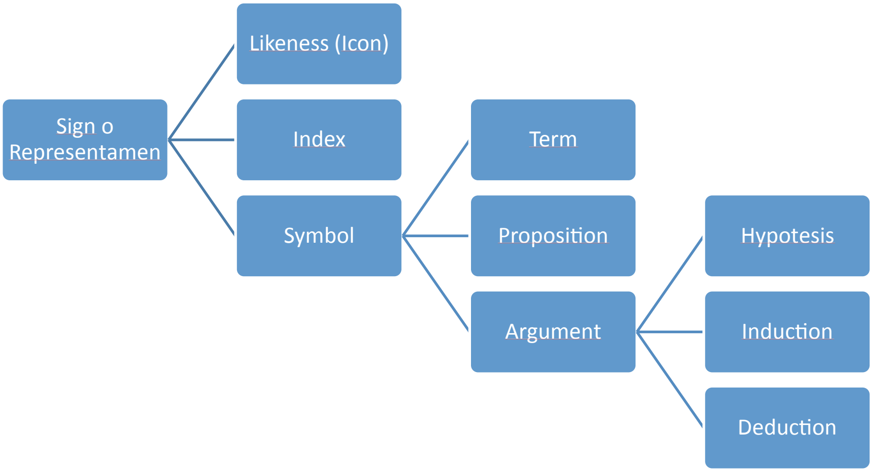
EP The Essential Peirce: Selected Philosophical Writings, edited by N. Houser and C. Kloesel, 2 vols. (Bloomington: Indiana University Press, 1992), followed by volume and page number.

L Reference to the correspondence of Peirce.

MS Peirce manuscripts in Houghton Library, followed by a Robin number and page number.

RLT *Reasoning and the Logic of Things*, edited by Kenneth Laine Ketner (Cambridge, Mass. Harvard University Press, 1992), followed by page number.

W Writings of Charles S. Peirce: A Chronological Edition, edited by Max Fisch et al. (Bloomington: Indiana University Press, 1982), followed by volume and page number.



As can be seen, the third element is always divided into three other elements. At that time, abduction was called Hypothesis, and none of the three types of arguments are further divided. An internal analysis of arguments made in his article 'Deduction, Induction and Hypothesis' from 1878 did not produce any classificatory output. There he proposed a syllogistic manner of analyzing arguments according to which the first proposition of a syllogism is a Rule, the second proposition is a Case, and the third proposition is a Result. Thus, a deduction is defined as a syllogism in which a case is subsumed in a rule to derive a result, induction is a syllogism that relates a case and a result to derive a rule, and a Hypothesis is a syllogism that derives a case by means a rule and a case. The analysis consisted in showing that the syllogistic form of induction and hypothesis corresponds to the inversion of the propositions of a deduction (see Table 1). The syllogistic account of logical inferences in his early years does not allow for any particular division within deduction, induction or abduction.

Thirty years later, Peirce abandoned the syllogistic approach to analyzing logical inferences (I will expose below the reasons he had in order to abandon the syllogistic analysis), and he took another path in order to analyze logical inference. The new criterion of analysis of logical inference is the role each of them plays in the scientific method. Abduction has a role of proposing a hypothesis, deduction has the role of tracking its consequences, and induction has the role of proving those consequence with experiments. Thus, the three logical inferences occupy an order of first, second and third. It is under this framework that Peirce began to state a

Table 1 Syllogistic account of logical inferences

	Deduction	Induction	Hipotesis
Major premise	Rule	Case	Rule
Minor premise	Case	Result	Result
Conclusion	Result	Rule	Case

clear classification of types of inferences in his article "On the Logic of Extracting the History of Ancient Documents" (1902), and continued to develop it over the years ending with his attempt in his article "A Neglected Argument for the Reality of God" (1908). These attempts at a classification follow the categorical principle of degeneracy by which all firstness is unique and simple; all secondness is a relation, a reaction, or a comparison, and, therefore, they take the forms of genuine and degenerate secondnesses; and all thirdness is a relation, unity, and generality, and, therefore, they take the forms of genuine and two degenerate thirdnesses (see EP2:160–61). From 1901 to 1908, each separate classification states that abduction is first, deduction is second, and induction is third. What varies between the different classifications is Peirce's establishment of subdivisions for deduction and induction. In consequence, abduction has not any kind; it is a unique and undivided kind of argument. Instead, deduction has two kinds: corollarial and theorematic; and induction has three kinds: crude, qualitative and quantitative. Editors of the Collected Papers realized that abduction has no subdivisions and said:

> His divisions, which reveal not only how his signs are classified, but why, for example, there is only one kind of abduction, two kinds of deduction, and three kinds of induction, are determined in part by the following principle: that which is a Second is divisible into two parts, of which one is itself divisible into two parts, etc., and that which is a Third is divisible into three parts, one of which is a First, another of which is a Second, (and thus divisible into two parts), and the last a Third, which is itself divisible into threes, and so on, apparently without end. (CP4, Introduction)

Similarly, Sami Paavola concludes that

> Peirce did not usually make any subdivisions in relation to abduction, and he seemed to be undecided if there are any. He even mentioned that he had found "no essential subdivision of Abductions" (Peirce PPM 276–277, 1903). A similar remark can be found at MS 764 (n.d.): "I have not been led to recognize any logical division of them [Retroductions]" (p. 31). [12].

Those are the reason Peirce considered abduction to have no divisions. However, as was mentioned, there are two passages in his late writings where he discusses that there are two types of abduction (not the same in both passages). The first passage corresponds to the final lecture he delivered in Harvard at 1903:

> The third cotary proposition is that abductive inference shades into perceptual judgment without any sharp line of demarcation between them; or, in other words, our first premises, the perceptual judgments, are to be regarded as an extreme case of abductive inferences, from which they differ in being absolutely beyond criticism. The abductive suggestion comes to us like a flash. It is an act of insight, although of extremely fallible insight. It is true that the different elements of the hypothesis were in our minds before, but it is the idea of putting together what we had never before dreamed of putting together which flashes the new suggestion before our contemplation.

On its side, the perceptive judgment is the result of a process, although of a process not sufficiently conscious to be controlled, or, to state it more truly, not controllable and therefore not fully conscious. If we were to subject this subconscious process to logical analysis, we should find that it terminated in what that analysis would represent as an abductive inference, resting on the result of a similar process which a similar logical analysis would represent to be terminated by a similar abductive inference, and so on ad infinitum (EP2, 227; CP. 5, 181).

It is clear that Peirce is establishing two types of abduction: that which is conscious, controlled, and open to criticism; and that which is unconscious, uncontrolled and beyond criticism. There is a continuous range of abductive inferences that goes from conscious to unconscious process "without any sharp line of demarcation between them". Consciousness and unconsciousness are the two limits of this continuous process; therefore, more than only two defined and determined points, abduction has an infinite range of possibilities and degrees of consciousness.

The second passage is an unpublished manuscript (MS 637) from 1909 that corresponds to his intended book on Meaning. He states that there are two types of abduction, namely, practical and scientific retroduction:

> The logical rules of inference from a surprising fact to a hypothetical state of things that would explain it involves a **complexity of conditions**, one of which is how much time can be allowed for coming to one's conclusion. A general who during a battle must instantly risk the existence of a nation either upon the truth of a certain hypothesis or else upon its falsity, must perforce go upon his judgment at the moment; and his doing so is in so far logical that all reasoning is based upon a tacit assumption that Nature in the sense of the aggregate of truth, is conformed more or less to something similar to the reasoner's Reason. This kind of inference may be called **Practical Retroduction**. It is ordinarily called a Presumption. If on the contrary, indefinite time can be allowed for judging of the truth of an explanatory hypothesis, the most favorable conclusion is that the consequences of that hypothesis that are susceptible of verification should be subjected to a systematic and thorough test. This which may be called **Scientific Retroduction** will be subject to considerations of economy. If, for example, a supposition far from reasonable, yet still possible, would explain a physical phenomenon, and if, in case it be incorrect, there is a way of disproving it at little expenditure of either time, energy, or any other valuable, then it may be worthwhile to clear the ground by taking up the investigation of it forth with. It will be remarked that the result of both Practical and Scientific Retroduction is to recommend a course of action. (MS 637, 1909; bold words added).

It is evident that Peirce distinguished between two kinds of abduction or retroduction, as it was called at that time: practical and scientific abduction. The criterion of distinction is the time available to resolve the doubt or surprising fact (see Niño, 201–202). The dual distinction only considers a long period of time or a short period of time, but between them, there can be a range of many possible differences. It is not the same the urgency that a doctor has in an ER for getting into the right diagnosis, that the urgency of deciphering a route to take into a complex transportation system in a city we are not familiarized with. Both are practical matters, and the time is not unlimited, but the consequences of erring in the hypothesis are different. Thus the spectrum of possibilities between a short and long period of time is infinite. In fact, following Peirce's synechism, it is possible to conceive that the difference between a practical and a scientific abduction is just a question of degree. Or where is the borderline that divides a practical and scientific inquiry?

Another important remark from this paragraph at MS 637 is that Peirce mentions that abduction depends on the "complexity of conditions". These conditions are the criteria according to which types of abduction can be derived. In this passage, he says that time is one of those conditions, but from his remark at Harvard Lecture, we can include also that consciousness is another condition. Moreover, it is possible to think of more conditions, and so, it is possible to have a complete theory of types of abduction. This will be discussed in the last part of this inquiry. For the moment, lets us focus on the opinions of other scholars on types of abduction and see what other criteria or complex conditions do they use in order to make those distinctions.

2 Contemporary Scholars on Types of Abduction

2.1 Two Types of Abduction

The first proposal of types of abduction made by contemporary scholars follows Peirce's remarks of two types of abduction. Thagard [18] was the first in considering the possibility of two kinds of abduction and called them *Hypothesis* or inference to a Case and *abduction* or inference to a Rule. Bonfantini and Proni [4] say it is a distinction that comes from Peirce himself who wrote about a low-level abduction close to sensation (they mean the perceptual judgment mentioned above), and high-level abduction that has the nature of argument which is close to scientific inquiry (see Bofantini and Proni [4, 129]). They add that the first type is involved in police investigations, for from a particular crime a particular cause can be inferred (1988: 126). The second type is involved in scientific inquiry, in which "the aim is to find a fundamental theoretical law of general application or (more often) to fit an anomalous fact into the applicability of a fundamental law by rearranging the "intermediary law". Although Eco rejects this distinction, he postulates it as follows: the first kind of abduction "starts from one or more surprising particular facts and ends at the hypothesis of another particular fact which is supposed to be the cause of the former (this seems to be the case of criminal detection)." Whereas the second type of abduction "starts from one or more surprising particular facts and ends at the hypothesis of a general law (this seems to be the case of all scientific discoveries)" [6, 204].

It seems that Thagard, Bofantini, and Proni confuse the two criteria already mentioned by Peirce in the two separated passages quoted above. They considered that the unconscious abduction represented in the perceptual judgments (low-level abduction) is the same type of the practical matters that lack enough available time, instantiated in a detective investigation. Perhaps they thought that the unconscious and uncontrolled type of abduction was due to the lack of time. But this is not the case; the perceptual judgments are unconscious and beyond criticism, because that is their nature. Although someone had plenty of time for observing or experiencing something, the perceptual judgment will continue to be unconscious. Those scholars did not realize they were facing not two types of abduction but perhaps four types.

One important thing they mention is that there can be two different types of abduction depending on the type of conclusion they reach, either a Case or a Rule. Although they confuse those two types of abduction with the other types proposed by Peirce, it is important to point out that this distinction can enrich the criteria of dividing types of abduction. If abduction concludes a case or a particular fact, they thought, this was the feature that makes it practical; and if an abductive inference concludes a Rule, that is, a law, this was the feature that makes it scientific. But remember that this was not Peirce's argument. The distinction between a practical and a scientific abduction depends upon the time available or the urgency with which one needs to reach a conclusion. Likewise, he never would have said that abduction can conclude Rules because that belongs exclusively to induction. Nonetheless, I underline that perhaps this distinction between a particular fact and a law can be one of those elements of the "complexity of conditions" that can derive types of abduction.

So far, then, we have three criteria that allow dual distinctions for abduction: consciousness, time, and types of conclusion. Let's keep this in mind for the last part of this inquiry.

2.2 Three and Four Types of Abduction

Despite the way that Bofantini and Proni proposed in the first part of their paper "To Guess or not to guess" that there is a dual division of abduction, they concluded at the end of the paper that there are three types of abduction that do not coincide with the two types above mentioned. They say:

> [It] is necessary to distinguish three principal types of abduction, with three ascending degrees of originality and creativity:
>
> ABDUCTION TYPE ONE –the mediation law to use for inferring the case from the result is given in an obligant and automatic or semiautomatic way.
>
> ABDUCTION TYPE TWO –the mediation law to use for inferring the case from the result is found by selection in the available encyclopedia;
>
> ABDUCTION TYPE THREE –the mediation law to use for inferring the case from the result is developed de novo, invented. It is in this last type of abduction that real guesswork comes in. (1988, 133–134).[3]

Umberto Eco [6] accepted Bofantini and Proni's suggestion of three types of abduction and added a fourth type. He named them as follows:

1. Hypothesis or overcoded abduction: The law is given automatically or semiautomatically.
2. Undercoded abduction. The rule must be selected from a series of equiparable rules put at our disposal by the current world knowledge.
3. Creative abduction. The law must be *invented ex novo*. To invent a law is not so difficult, provided our mind is "creative" enough.

[3]In a posterior article, Bofantini extended this proposal, particularly the third type, (Bonfantini, 1993: 326–327; cf. Bonfantini, 1987: 69, 112; Niño, p. 265), which is divided into three subtypes. Nonetheless, since the whole proposal is going to be criticized below, there is no need to enter into details of these further distinctions, for they also fall into the general critique.

4. Meta-abduction. It consists in deciding as to whether the possible universe outlined by our first-level abductions is the same as the universe of our experience. (see Eco [6, 210–220])

As can be seen, Bofantini's, Proni's and Eco's division of three first types of abduction is based on the syllogistic analysis of inferences as composed by a rule, a case, and a result. According to it, a Hypothesis is a syllogism that derives a case by means a rule and a case. Bofantini's, Proni's and Eco's division of the three first types of abduction focusses in the Rule (one of the premises), law or "mediating law" as they call it, and attempt to determine its nature. Hence, they claim that, in the first type of abduction, the rule or law is obligant or given automatically; in the second type, the law is selected from the "available encyclopedia" or "current world knowledge"; and in the third type, the law is invented.[4]

Some difficulties are presented in this triadic distinction based on the nature of the law. First, how does it establish the three different types of rules or laws? Is this distinction sound and consistent? The first type of law is obligant. Are not all laws obligant? Is it not the nature and definition of law to obligate all particular cases subsumed in it? Whether the law is selected or invented (second and third type) is a law and it must be obeyed by all particular cases covered by it. The second type of law is selected from the "available encyclopedia", namely, it is a law that we already know. Is this different to the first type of laws (obligants)? This seems to be a false dichotomy; an obligant law can be already known, and all known laws have to be obligant. Therefore, the first and second types of abduction do not establish any real distinction.

The third type of law is invented. This is the most complex question that this distinction raises, for it creates, from my point of view, a circular argument. It says that creative abduction is made after a law is invented. Is not the purpose of abduction to create or conjecture laws or conceptions? Is not also a purpose of abduction to take us from lack of knowledge to knowledge? If a law must be invented before abduction can take place, then a previous abduction must be made, and so on. What kind of abduction should be done before 'creative abduction' takes place? Therefore, 'creative abduction' is a *petitio principium* and its application makes a regression ad infinitum.

Secondly, another difficulty that this distinction presents is that it focuses only in the rule or law. What happens with the other premise (result) and with the conclusion

[4]A similar distinction between the second and third type of abduction is made by Campos [5] and Magnani [10] (see also, [17, 202]). Campos states that there are two types of abduction, namely: a 'habitual abduction' and a 'creative abduction'. Magnani [10] distinguishes between 'selective abduction' that chooses the best hypothesis among a manifold of hypothesis already known, and 'creative abduction' that introduces new concepts and laws. They are similar to Eco's 'undercoded abduction' and creative abduction' because they focus on the nature of the second premise, rule or law. Campos' 'habitual abduction' and Magnani's 'selective abduction' are that inference that results when the law is already known and the surprising fact is explained under that law. 'Creative abduction' is that in which the law is unknown, and, therefore, a law must be conjectured before proceeding to explain the surprising facts [5]. Consequently, the critiques that I am about to present on Eco's distinction include also Campo's and Magnani's distinction.

(case)? Would not be this other premise also obligant, selected and invented? If this were the case, then it would not be three types of abduction but twenty-seven (3^3), depending on the nature of its propositions (either premises or conclusion).

Thirdly, the classification of the law as obligant, selected and invented imitates the modal distinction between: possible, actual, and necessary. In the case of Peirce, it refers to his theory of categories according to which firstness is possible, secondness is actual, and thirdness is necessary. Hence Bofantini's, Proni's and Eco's distinction is upside down; the first type of abduction should be the invented one, for it is mere possibility; the second type is the selected one for it is an actual law; and the third type of abduction should be necessitating or obligant as they call it.

The first objection is sufficient to consider that those distinctions (obligant, selected, and invented) do not have any consequence in establishing types of abduction. Regardless of the origin or nature of the law, the abduction that uses such a law as a premise does not have any change. The nature of inference depends on the habit according to which it passes from the premises to the conclusion. The nature or the origin of the premises does not modify the nature of the inference. A deduction, for instance, can have false premises, included a false law, such as "all men are flying beings", and still conclude necessarily and validly that "Socrates is a flying being". Thus, abduction can have whatever premises one prefers (obligant, selected, or invented) and still it can suggest or conjecture that an event belongs to a rule.

In consequence, I reject entirely this procedure to determine types of abduction.[5] It is based on a triadic distinction that is not sound; it lacks completeness for it only takes into account the nature of the Rule (second premise); it lacks precision in determining the order in which those distinctions should be made; and therefore it lacks coherence with Peirce's categories.

The fourth type of abduction by Eco, or meta-abduction, should be rejected as well for it surpasses the frontier of abductive reasoning. It takes the place of the verifying step of inquiry that Peirce identifies with induction. To decide if abduction is the same as the universe of experience is the same as verifying or proving. The contact with experience belongs only to induction, for induction is the inference asking nature if the hypothesis is true or not. Obviously, this verifying step is indispensable in every inquiry, and every abduction must be tested, but its verification corresponds to induction. In conclusion, Eco's division of four types of abduction should not be accepted.

Thagard [19] also claims there are four kinds of abduction, but unlike Eco's classification, this division depends on the nature of the conclusion of the abductive inference. These four kinds are called simple, existential, rule-forming, and analogical abductions. He defines them as follows:

> Simple abduction produces hypotheses about individual objects (…). Existential abduction postulates the existence of previously unknown objects, such as new planets. Rule-forming

[5]Paavola is also a critique of this procedure. He says: "I think it is not, for example, clear in which sense "creative" abduction can start totally "ex novo", and if overcoded abduction is more inductive than abductive, and if meta-abduction is a distinct type of abduction (or rather a second-order abduction)." (2006, pp. 49–50).

> abduction produces rules that explain other rules and hence is important for generating theories that explain laws. Finally, analogical abduction uses past cases of hypothesis formation to generate hypotheses similar to existing ones. [19]

This new classification by Thagard [19] is an extension of his classification of two types of abduction commented *supra* [18]. There he considered one type as an inference to a Case and a second type as an inference to a Rule, that is, there were two types of abduction depending upon the types of conclusions they reach. Now, this new classification enlarges this same distinction and claims that the particular facts concluded through abduction can be either known and observable (simple abduction) or unknown and unobservable (existential abduction). For instance, if you see a group of young men outrageously dressed, then you may conclude they are rock musicians, based on the rule that all rock musicians dress outrageously; this is a simple abduction. But if an astronomer sees that the orbit of Uranus is perturbed, it may be concluded that there is another planet that is unobserved which happens to be Neptune; this is an existential abduction (see Thagard [19]).

The third type proposed by Thagard is that which concludes rules, theories, or general laws. It cannot be a simple generalization in which a rule like all A are B is postulated in order to explain why a particular A is B. "The first is problematic and does not seem to play any role in theory formation. In it, we postulate the rule that all A are B to explain why a particular A is B. The second is a combination of abduction and generalization" [19]. Hence, rule-forming abduction should not be confused with generalization or induction. For example, the next syllogism concludes a universal statement, but it is not a rule-forming abduction: Michael dresses outrageously; Michael is a rock musician; therefore, all rock musicians dress outrageously. This is clearly an induction. It is a syllogism in the third figure that concludes a general proposition. It is precisely what Aristotle indicated as induction (Prior Analytics, II, 23) and it is exactly what Peirce called a perfect form of induction (see W1: 177, 179, 262–264; and MS 741, p. 11 ff.; Cf. Flórez [8, 42–50]).

The abduction to rules that Thagard actually accepts is that in which a generalization is mixed with an explanation. "Abduction to rules works like: if PI forms the abduction that the F that is x is also H because that would explain why is G, it can naturally generalize that all F are H." (1988, 59). The difference for induction and simple generalization is that the syllogism of an abduction to rule also includes the explanation of the facts exposed in the premises, not just their generalization. Hence, abduction to a rule and induction has the same syllogistic form, but, as Thagard recalls it, "not all induction is an inference to an explanation" (Ibid). Thagard's example of abduction to rules is the inference to the law "all sounds are waves" based on the characteristics of sound that will be explained if sound is a wave.

The fourth type of abduction by Thagard is analogical abduction. It consists in postulating hypothesis similar to other satisfactory and successful hypotheses. "Once a theory has established itself, scientists frequently want to use similar kinds of explanations." (1988, 60). For instance, if Sherlock Homes has already resolved cases with the hypothesis that a criminal had a particular motive, it is plausible to conjecture that a new crime was committed also with a particular motive. Another example is "how the great success of Newtonian mechanics made most scientists of the eighteenth and nineteenth centuries strive to give mechanical explanations of phenomena" (Ibid). Similarly, Thagard claims that apparently Darwin's hypothesis of evolution of species was based on the analogy with the artificial selection by breeding (Cf. [19]).

Thagard's classification of 1988 is sounder than Eco's but it is incomplete for it is based only on the types of conclusions that abductive inferences render; it does not pay attention to the evidence or premises on which they are based. Moreover, the fourth type or analogical abduction does not correspond to the same criteria with which the three first types of abduction were distinguished. Analogical abduction does not pay attention to the types of conclusion but to the guiding principle that rules abduction. Notice that any of the three first types of abduction can be analogical. So, it is necessary to identify exactly what is the criterion that determines this classification and to identify some other types of inferences that this criterion permits. Consequently, Thagard's criteria will be considered in the final part of this inquiry but they should be extended and complemented.

2.3 Eleven Types of Abduction

The extension of these criteria was precisely proposed by Schurz [17]. Threefold and fourfold divisions of abduction tried so far were based separately either in the nature of the intermediary law (second premise or rule) or in the nature of the conclusion reached. Schurz [17] proposes a more extensive classification based on these two criteria already expressed by other scholars, plus a third criterion or dimension, as he calls them:

1. The kind of premises or evidence available to make abductive inferences. The alternatives here are two: there are either facts or laws.
2. The kind of conclusion that is abducted. Here there are basically three alternatives: the conclusion is either a fact or a law, and the facts can also be divided into observable facts or unobservable facts.
3. "The beliefs or cognitive mechanisms which drive the abduction." The alternatives here are three: the abduction is driven by law, scientific theories, or causal unification (see 2008, 205).

If these options are combined, there appear four genera of abduction: 1. Factual abduction; 2. Law abduction; 3. Theoretical model abduction, and 4. Second order existential abduction. Some of these genera are subdivided; for instance, factual abduction has three species: (a) observable fact abduction, (b) First order existential abduction, and (c) Unobservable fact abduction. Instead, law abduction and theoretical model abduction has no subdivisions. Finally, the fourth genus has more subdivisions that the others. The next chart shows all those types of abduction. The final subdivision is that which constitutes a definite type of abduction (their name are in italics), and therefore, Schurz's classification can be summarized in eleven types of abduction.

The first genus is factual abduction. It is defined by Schurz as that inference in which "both the evidence to be explained and the abducted hypothesis are singular facts. Factual abductions are always driven by known implicational laws going from causes to effects, and the abducted hypotheses are found by backward reasoning, inverse to the direction of the lawlike implications." (2008, 206) The best examples of this type of abduction are detective investigations, for a crime scene is a fact and the conjectured hypothesis is also a fact. For sake of brevity, I am not going too deeply into the subclasses of factual abduction (see [17]). As can be seen in Table 2, all factual abductions coincide in their general structure.

The second genus is law abduction. As its name indicates, this type of abduction is that which is supported and driven by law, and more importantly, it concludes laws. Given two general statement such as 'whatever contains sugar is sweet' and 'all pineapple is sweet', it is possible to conjecture the law: "All pineapple contains sugar".[6]

The third genus is theoretical model of abduction. Schurz explains it as: "The explanandum of a theoretical-model abduction is typically a well-confirmed and reproducible empirical phenomenon expressed by an empirical law (…). The abduction is driven by an already established scientific theory which is usually quantitatively formulated. The abductive task consists in finding theoretical (initial and boundary) conditions which describe the causes of the phenomenon in the theoretical language and which allow the mathematical derivation of the phenomenon from the theory" (2008, 213). So the difference between law abduction and theoretical model abduction lies on the difference between law and theory. Schurz explains theoretical model abduction as the work of normal science, in Kuhn's terms (2008, 213). A theory is explained by Schurz as the work on a normal science (Kuhn) in the sense in which a normal science always want to explain an anomaly in terms of the dominant theory.

The fourth genus is called second order existential abduction. It is the one that has more subdivisions; they depend mostly upon the mechanism that drives the inferences, either extrapolation, analogy, or, pure unification. Thus, the first species of second order existential abduction is micro part abduction, and it is defined as the extrapolation of the knowledge one has on the macroscopic world and it is conjectured that the microscopic world is similar. The best example of this type of hypothesis is precisely the atomistic world view. The second subspecies of second order existential abduction is analogical abduction. As its name indicates, this is an abduction driven by analogy. It is the same analogical abduction proposed by Thagard [19]. The next subspecies of second-order existential abduction is hypothetical cause abduction. It has two types; the first one conjecture the existence of an unobservable cause and it is driven by speculation (no particular law, theory, or background); the second type

[6]According to Schurz, this type of abduction is what Aristotle had in mind in Posterior Analytics I, 34. This passage is what I consider is the true origin of abduction in Aristotle (see Flórez [7, 272–277]).

Table 2 Schurz's classification of eleven kinds of abduction (see [17])

Kind of abduction				Evidence to be explained	Product of abduction	Abduction is driven by
Factual abduction or retroduction	*Observable-fact abduction*			Facts	Observable facts	Known laws
	1st order existential abduction			Facts	Observable facts	Known laws
	Unobservable Fact abduction (historical abduction)			Facts	Unobservable facts	Known laws
Law abduction				Empirical laws	New laws	Known laws
Theoretical model abduction				General empirical laws	New theoretical models	Known theories
2nd order existential abduction	*Micro-part abduction*			Laws	New laws or new theories with new concepts/microscopic composition	Theoretical background knowledge/extrapolation of background and knowledge
	Analogical abduction			Laws	New laws or new theories with analogical concepts	Analogy with background knowledge
	Hypothetical Cause abduction	*Speculative abduction*		Laws	Hidden unobservable causes	Speculation
		Common cause abduction	*Strict common cause abduction*	Laws	Hidden common causes/new theoretical concepts	Causal unification
			Statistical factor abduction	Laws	Hidden common causes/new theoretical concepts	Causal unification
			Abduction to reality	Introspective phenomena	Hidden common causes/concepts of external reality	Causal unification

is conjectures a common cause and it is driven by the search of causal unification (it does not require any background knowledge). The latter has some other subdivisions that I am not going to expose for the sake of brevity (see [17]).

In conclusion, Schurz's classification presents a series of criteria that allows classifying types of abduction. The criterion of the product of abduction, either a fact or a law, was already considered by Thagard [18, 19], although Schurz includes the variable of the fact being observable or unobservable. The other two criteria are novel and they seem very promising. To consider the type of premises or evidence available to make the abductive inference is important. Bofantini, Proni, and Eco based their classification on the types of laws included in the premises; but they neglect the possibility of an abduction based entirely in particular facts. The last criterion is the most novel one, for it includes "The beliefs or cognitive mechanisms which drive the abduction." It seems that it refers to the 'guiding principle' that Peirce claims to drive every inference. He says:

> The particular habit of mind which governs this or that inference may be formulated in a proposition whose truth depends on the validity of the inferences which the habit determines, and such a formula is called a guiding principle of inference. Suppose, for example, that we observe that a rotating disk of copper quickly comes to rest when placed between the poles of a magnet, and we infer that this will happen with every disk of copper. The guiding principle is, that what is true of one piece of copper is true of another. Such a guiding principle with regard to copper would be much safer than with regard to many other substances – brass, for example. (EP1, 112)

Peirce did not mention any particular kind of guiding principle, although he recognizes there can be plenty of them: "A book might be written to signalize all the most important of these guiding principles of reasoning." Schurz recognizes six of those habits of mind or guiding principles: a law, scientific theories, background knowledge, analogy, speculation, and causal unification.

2.4 Fifteen Types of Abduction

Hoffman [9] works on Schurz's classification enlarging and reclassifying other types of abduction. Hoffmann's classification reaches a total of fifteen types of abduction. He criticizes Schurz's criteria to determine types of abduction for they "simply do not allow for one coherent list" (2010, 579). Hoffmann's classification is based only on the following two criteria:

1. There are five things that can be abductively inferred: facts (observable or unobservable), types, laws, theoretical models, and systems of representation. [9, 579, 584]
2. There are three different procedures by which abduction can be made: first, selective abduction according to which an explanatory hypothesis from a given database; Second, personal creative abduction according to which a new explanatory hypothesis is made (P-creative); and third a collective creative abduction

Table 3 Hoffman's fifteen kinds of abduction (*Source* Hoffman [9, 585])

	Selective abduction: hypothesis already exists in our minds	P-creative: hypothesis is new for us, but it exists already in culture	H-creative: hypothesis is historically new
The explanatory hypothesis is a new fact	Selective fact abduction	P-creative fact abduction	H-creative fact abduction
The explanatory hypothesis is a new type or concept	Selective type abduction	P-creative type abduction	H-creative type abduction
The explanatory hypothesis is a law	Selective law abduction	P-creative law abduction	H-creative law abduction
The explanatory hypothesis is a theoretical model	Selective model abduction	P-creative model abduction	H-creative model abduction
The explanatory hypothesis is based on a new system of representation	Selective meta-diagrammatic abduction	P-creative metadiagrammatic abduction	H-creative metadiagrammatic abduction

according to which a new explanatory hypothesis is made that no one else has made in history (H-creative).

If these two criteria are combined, fifteen types of abduction result (five things to be inferred multiplied for three procedures). Consequently, Hoffman adds a fifth genus of abduction to the four main genera of abduction proposed by Schurz. His new genus is called "meta-diagrammatic abduction" (2010, 574, 579–584). Every genus is defined by the reference to the first criterion of things inferred: factual abduction, type abduction, law abduction, theoretical model abduction, and meta-diagrammatic abduction. Moreover, every genus has three species according to the three procedures: selective, P-creative and H-creative (Table 3).

Although Hoffman offers a more comprehensive classification of types of abduction than that of Schurz, he neglects the criterion of the types of premises. He focuses only on the explanatory hypothesis (conclusions) and the guiding principles that drive inferences. This last criterion is even diminished with respect to the progress made by Schurz who proposed six guiding principles. Hoffman's distinction among selective, P-creative and H-creative suggests Bofantini, Proni and Eco's distinction among obligant, selective and creative abduction. It does not really focus on guiding principles, for it focuses on the origin of the intermediary law that rules the inference. The origin can be from either a given data (selective) or a new law for the inquirer (P-creative) or something completely new for everyone (H-creative).

Finally, the last proposal by a contemporary scholar that will be discussed is Aliseda [1], who despite the fact does that she not propose a classification of types of abduction, she indeed enlists three different parameters, as she calls them, which will allow a further classification and can be compared with the other criteria proposed by

Peirce and other scholars. “Our aim is not to classify abductive processes, but rather to point out that several kinds of these are used for different combinations of the above parameters” (2016, 48). Those parameters are, first, the inferential parameter that “sets some suitable logical relationship among *explananda*, background theory, and *explanandum”* (2016, p. 46). This parameter is similar to the guiding principle or principle that drives the abductive inference. She claims that the inference parameter “ranges over such diverse values as probable inference ($\Theta, \alpha \Rightarrow$ probable ϕ), in which the *explanans* renders the *explanandum* only highly probable, or as the inferential mechanism of logic programming ($\Theta, \alpha \Rightarrow$ prolog ϕ).” (2016, 46)

The second parameter is the kind of phenomenon (φ) that triggers the abductive inference. Namely, it refers to the starting point that forces the abductive inferences. Aliseda distinguishes between two types of triggers: novelty and anomaly. Peirce had thought that the only trigger of abduction is a surprising fact. Aliseda, instead, thinks that surprise is a relative concept for it depends on the background knowledge everyone has. Hence, in order to be more precise, she distinguishes between two types of surprising facts: novelty is a surprising fact that is consistent with theory, whereas anomaly is a surprising fact that is inconsistent with theory.

Abductive Novelty: $\Theta \Rightarrow \phi, \Theta \Rightarrow \neg\phi$

ϕ is novel. It cannot be explained ($\Theta \Rightarrow \phi$), but it is consistent with the theory ($\Theta \Rightarrow \neg\phi$).

Abductive Anomaly: $\Theta \Rightarrow \phi, \Theta \Rightarrow \neg\phi$.

ϕ is anomalous. The theory explains rather its negation ($\Theta \Rightarrow \neg\phi$). (2016, 47).

The third parameter is the abduction product or outcome; in other words, it is the type of conclusion or explanatory hypothesis that is derived from abduction. She recognizes three types of outcome: facts, rules, or theories.

Aliseda did not commit to a classification of types of abduction; but it is easy to see that if those three parameters were combined, there would be eighteen types of abduction: three inferential parameters multiplied by two triggers give as a result six types of abduction, and those six types multiplied by three types of outcomes give as a result eighteen types of abduction.

Conclusively, the journey across proposals of types of abduction (see Table 4) has allowed us to identify different criteria to distinguish kinds of abduction. Peirce himself knew there was a “complexity of conditions” (MSS 637) to determine abductive inferences, and he identified only two of those conditions: first, available time or urgency to reach a conclusion with which he distinguishes between a practical and a scientific retroduction; and second, consciousness and the possibility of criticism with which an abduction is made, so he distinguished between perceptual judgments and abduction. Thagard recognizes in the types of conclusion a new criterion of classification; first he postulates only two types: either an abduction concludes cases or particular facts, or it concludes rules or laws (1978). Later he postulates that particular facts can be either observable or unobservable, that abduction can conclude laws, and that abduction can be made analogically (1988). Bofantini, Proni, and Eco focus their analysis in another criterion: the nature of the intermediary law, second premise

Table 4 Classification of types of abduction according to their affinity

Peirce (1903)	Bofantini and Proni 1	Peirce (1909)	Bofantini and Proni 2	Eco	Campos [5]	Magnani [10]	Thagard [18]	Thagard [19]	Schurz [17]		Hoffman [9]	
Perceptual judgment	Low-level abduction	Practical retroduction	Abduction type 1: obligant law	overcoded abduction: obligant law			Hypothesis or abduction to a case	Simple abduction	Observable fact abduction	Factual abduction	Selective fact l abduction	Factual abduction
			Abduction type 2: selected law	Undercoded abduction: selected law	Habitual abduction	Selective abduction			1st order existential abduction		P-creative fact abduction	
			Abduction type 3: invented law	Creative abduction: invented law	Creative abduction	Creative abduction		Existential abduction	Unobservable fact abduction (historical abduction)		H-creative fact abduction	
				Meta abduction: decision whether the conjecture corresponds to experience			Abduction or abduction to a rule	Rule-forming abduction	Law abduction		Selective type abduction	Type abduction
Abduction	Scientific abduction	Scientific retroduction							Theoretical model abduction		P-creative type abduction	
									Micro-part abduction	Second order existential abduction	H-creative type abduction	
								Analogical abduction	Analogical abduction		Selective law abduction	Law abduction
									Speculative abduction		P-creative law abduction	

(continued)

Table 4 (continued)

Peirce (1903)	Bofantini and Proni 1	Peirce (1909)	Bofantini and Proni 2	Eco	Campos [5]	Magnani [10]	Thagard [18]	Thagard [19]	Schurz [17]		Hoffman [9]	
									Strict common cause abduction		H-creative law abduction	
									Statistical factor abduction		Selective model abduction	Theoretical model abduction
									Abduction to reality		P-creative model abduction	
											H-creative model abduction	
											Selective meta diagrammatic abduction	Meta-diagramamatic abduction.
											P-creative meta-diagrammatic abduction	
											H-creative meta-diagrammatic abduction	

or rule on which every abduction is based. Thus, they postulate that laws can be obligant, selective, or invented [4, 133–134; 6, 206–207]. Moreover, Eco added a fourth type whose criterion is not the nature of the law, but the possibility of coincidence of the conclusion with experience. Aliseda [1] claims there are three parameters: inferential parameters (probabilistic or programming), triggers (novelty or anomaly) and outcome (facts, laws or scientific theories). Schurz [17] proposes three criteria some of which were already proposed; firstly, he considers the kind of premises abduction possesses, either facts or laws; secondly, he considers the kind of conclusion abduction reaches, observable facts, unobservable facts or laws; and thirdly, he considers the cognitive mechanism that drives abduction: law, scientific theories or causal unification. Finally, Hoffmann [9] recognizes two criteria of distinction: firstly, the conclusion of explanatory hypothesis abduction reaches facts (observable or unobservable), types, laws, theoretical models, and systems of representation. Secondly, procedures by which abduction can be made; he identifies three kinds of procedures: from a given database; from a new personal explanatory hypothesis (P-creative); and from a new historical explanatory hypothesis (H-creative).

3 A Conjecture of Types of Abduction Based on Previous Criteria

Now, a comprehensive classification of types of abduction should take into account many of these criteria, except for those that were already rejected in the previous section. A way of organizing these criteria is through a distinction between external and internal conditions of abduction. Thus, internal conditions are those that belong to the syllogistic or argumentative form of the abductive inference such as the nature of the premises and the nature of the conclusions. External conditions are those circumstantial criteria such as time, consciousness, triggers, and guiding principles. Before entering to analyze internal and external conditions of abductive inferences, it is necessary to say that with this distinction there can be made a possible response to Peirce's paradox according to which types of abduction are postulated but it is asserted several times that abduction has one only type. It seems that Peirce thought that the internal conditions of an abductive inference are the same; in his early account he considered that the formal structure of abduction is a syllogism in the second figure that attempts to infer a Case, and in his mature account is always a hypothetical syllogism of the form:

The surprising fact, C, is observed;
But if A were true, C would be a matter of course,
Hence, there is a reason to suspect that A is true.[7]

Therefore, he never considered that the internal conditions of an abductive inference were relevant; but the external condition of time and consciousness can modify

[7]EP2: 231; CP 5. 189.

the abductive inference. In other words, for Peirce, a practical or scientific retroduction always have the same structure, the difference is only how much time there is available to reach a conclusion; or a perceptual judgment or an abduction properly also has the same structure (if they were formalized) but the only difference is that one is controllable and made consciously, but the other is uncontrollable. Consequently, it seems that when Peirce advocates for a unique type of abduction from a categorical point of view, he is referring to the internal conditions. Whereas when he advocates for two types of abduction he is referring to the external conditions.

The internal conditions are referred entirely to the syllogistic form of inferences. In his early account of inference, Peirce took this path but he never considered types of abduction; moreover, later on his career, he considered the syllogistic account of inferences as not relevant. He said: "I was too much taken up in considering syllogistic forms and the doctrine of logical extension and comprehension, both of which I made more fundamental than they really are." (CP2. 102 from 1902). Likewise, he said that "I do not think, however, that this formal way of defining Induction and Abduction by reference to the syllogism conveys much idea of their real nature except to a mind that has penetrated very deeply into the nature of syllogism" (MS 475, ISP12-17, 1903). Peirce did not reject this kind of analysis at all but recognized it to be not as fundamental and important.[8] At this respect, says Niño: "On the other hand, it is important to point out that Peirce just says that that 'formal presentation' of abduction and induction is not 'very communicative' for someone no familiarized with syllogisms. He does not say that it is erroneous, confusing, refuted or overcome to understand them in that way. Moreover, he says, once again, that forging a hypothesis to explain observed facts is the inference of the minor premise by means of the major and the conclusion. In other words, abduction is the inference of an antecedent through a consequence (rule) and a consequence (result)" [11, p. 175. Translation is mine].[9]

This, once again, can be the reason why Peirce only considered the external conditions of time and consciousness but neglected the internal conditions of the syllogistic account. The external conditions make him think that there are two types of abduction, and the neglect of internal conditions make him think there are no types of abduction. And so, the paradox with which this inquiry started is solved (at least for me the doubt ceased).

Now, if we, despite Peirce's neglect, pay attention to the internal condition of the syllogistic account, then we can include the criteria proposed by other scholars. In particular, we can synthesize those proposals in two main criteria:

---

[8]Belluci (2018, 2) consider instead that the syllogistic account of inferences was incorrect whatsoever.

[9]Por otra parte, hay que notar que Peirce sólo dice que esa 'presentación formal' de la Abducción y de la Inducción no es 'muy comunicativa' para alguien no familiarizado con el silogismo. No dice que sea erróneo, confuso, refutado o superado entenderlo así. Además, dice, una vez más, que el forjar una hipótesis para explicar hechos observados es la inferencia de la premisa menor a partir de la premisa mayor y la conclusión. Es decir, la Abducción es la inferencia de un antecedente (Caso) a partir de una consecuencia (Regla) y un consecuente (Resultado).

1. The types of premises or evidence on which an abductive inference is based: there can be particular facts or laws.
2. The type of conclusions or explanatory hypothesis that results from the abductive inference: there can be also particular facts or laws.

These two criteria, as you can see, are related to the quantity and quality of propositions and that precisely is what the syllogistic account of logical inferences does. Peirce's presentation of the syllogistic account (1878) did not explore entirely the possibilities of different types of propositions. I suggest that an extension of the syllogistic account can convey types of abduction based on distinctions on types of premises and types of conclusion.

As was already mentioned, the syllogistic account of inferences states a general form to abduction (hypothesis) as a syllogism that derives a Case by means a Rule and a Result. It is so because it takes each proposition of the syllogism and regardless the quantity or quality of those propositions, it assigns them a function: either a rule, a case, or a result. But if all those characteristics are taken into account, then the same number of valid syllogism that Aristotle and the medieval thinkers found for deduction can be also assigned to induction and abduction but with the propositions reversed or relocated. In the general description of the three types of arguments, Peirce stated that deduction is identified as Barbara, induction as Baroco and abduction as Bocardo (see EP1, 190). But in a specific determination of every form of the proposition, then deduction has twenty-four types (Aristotelian and Theophrastean syllogisms, such as Barbara, Celarent, Darii, Ferio, etc.). It seems valid to me that the same sort of deductive syllogisms make the same number of inductions and abductions with the proposition reversed. And since induction is a reversed deduction that tries to infer the rule or major premise from a case and a result, and since abduction is a reversed deduction that infers a case by means of a result and a rule, then there are as many valid types of abductions and inductions as there are valid moods and figures of deductive syllogism.

Peirce's example on his article "Deduction, Induction, and Hypothesis" from 1878, starts with a deductive syllogism in Barbara and reversing its proposition obtains its corresponding induction and abduction (Table 5).

Table 5 Symmetry of the three kinds of inferences (Charts constructed from the article "Deduction, Induction, and Hypothesis" (1878), EP1: 188)

	Deduction	Induction	Abduction
Major premise	All the beans from this bag are white (rule)	These beans are from this bag (case)	All the beans from this bag are white (rule)
Minor premise	These beans are from this bag (case)	These beans are white (result)	These beans are white (result)
Conclusion	These beans are white (result)	All the beans from this bag are white (rule)	These beans are from this bag (case)

This procedure executed by Peirce is instantiated only by a deductive syllogism of Barbara (Mood AAA, and figure 1), and it derives an inductive syllogism AAA-3, and it derives an abductive syllogism AAA-2. This procedure indicates that it is possible not only to reverse the propositions of a syllogism Barbara. My proposal is that it is possible to extend it to all twenty-four valid deductive syllogisms accepted by tradition. For the sake of brevity, I limit the following exercise to abduction, but it is clear that it can also be extended to induction (Table 6).

Table 6 Types of abduction according to the reversal of the same propositions of deduction

Deductive syllogisms			Abductive syllogisms		
First figure	Barbara	All M is P (rule) All S is M (case) All S is P (result)	Second figure	All M is P (rule) All S is P (result) All S is M (case)	AAA-2
	Barbari	All M is P All S is M Some S is P		All M is P Some S is P All S is M	AIA-2
	Celarent	None M is P All S is M None S is P		None M is P None S is P All S is M	EEA-2
	Celaront	None M is P All S is M Some S is not P		None M is P Some S is not P All S is M	EOA-2
	Darii	All M is P Some S is M Some S is P		All M is P Some S is P Some S is M	AII-2
	Ferio	None M is P Some S is M Some S is not P		None M is P Some S is not P Some S is M	EOI-2
Second figure	Baroco	All P is M Some S is not M Some S is not P	First figure	All P is M Some S is not P Some S is not M	AOO-1
	Cesare	None P is M All S is M None S is P		None P is M None S is P All S is M	EEA-1
	Cesaro	None P is M All S is M Some S is not P		None P is M Some S is not P All S is M	EOA-1
	Camestres	All P is M None S is M None S is P		All P is M None S is P None S is M	AEE-1
	Camestros	All P is M None S is M Some S is not P		All P is M Some S is not P None S is M	AOE-1

(continued)

Table 6 (continued)

Deductive syllogisms			Abductive syllogisms		
	Festino	None P is M Some S is M Some S is not P		None P is M Some S is not P Some S is M	EOI-1
Third figure	Bocardo	Some M is not P (rule) All M is S Some S is not P	Second figure	Some M is not P Some S is not P All M is S	OOA-2[a]
	Darapti	All M is P All M is S Some S is P		All M is P Some S is P All M is S	IAA-2*
	Datisi	All M is P Some M is S Some S is P		All M is P Some S is P Some M is S	IAI-2*
	Disamis	Some M is P All M is S Some S is P		Some M is P Some S is P All M is S	IIA-2*
	Felapton	None M is P All M is S Some S is not P		None M is P Some S is not P All M is S	OEA-2
	Ferison	None M is P Some M is S Some S is not P		None M is P Some S is not P Some M is S	OEI-2*
Fourth figure	Bamalip	All P is M All M is S Some S is P	Fourth figure	All P is M Some S is P All M is S	IAA-4
	Calemes	All P is M None M is S None S is P		All P is M None S is P None M is S	EAE-4*
	Calemos	All P is M None M is S Some S is not P		All P is M Some S is not P None M is S	OAE-4*
	Dimatis	Some P is M All M is S Some S is P		Some P is M Some S is P All M is S	IIA-4*
	Fesapo	None P is M All M is S Some S is not P		None P is M Some S is not P All M is S	OEA-4*
	Fresison	None P is M Some M is S Some S is not P		None P is M Some S is not P Some M is S	OEI-4*

[a]The figure of syllogisms depends mainly on the place of the middle term. Nonetheless, the minor and major premises depends on the place on the position of the minor and major terms. The minor term is the subject of the conclusion and the major is the predicate of the conclusion. In order to comply with these rules, the syllogisms identified with an asterisk (*) modify the order of the premises

Table 7 Thirty-two abductive syllogisms according to types of premisses and types of conclusions

	Explanatory hypothesis is a particular fact	Explanatory hypothesis is a law
Premises are particular facts	*OOI-2*; *III-2*; *III-4*	OOA-2; IIA-2; IIA-4
Premises are facts and laws	AII-2; EOI-2; AOO-1; EOI-1; IAI-2; OEI-2; OEI-4	AIA-2; EOA-2; AOE-1; EOA-1; IAA-2; OEA-2; IAA-4; OAE 4; OEA-4.
Premises are laws	*AAI-2*; *EEI-2*; *EEI-1*; *AEO-1*; *EAO-4*	AAA-2; EEA-2; EEA-1; AEE-1; EAE-4

Accordingly, there are twenty-four abductive syllogisms. The two first criteria can be partially explained with the syllogistic account as was just extended. The twenty-four different abductive syllogisms show that there can be both, either particular or universal premises, and, either particular or universal conclusions. But they do not fill all the possibilities. The next chart (Table 7) shows that these twenty-four syllogisms fill four options and leave blank two other options. However, the empty options can easily and necessarily be filled by means of a deductive principle according to which a particular follows from a general, as a replica, Peirce would say. Therefore, if from particular premises it is possible to conjecture a hypothesis that has the form of a universal proposition, then it is possible also to conjecture a particular hypothesis. In the same way that from Barbara is possible to deduce Barbari, thus in an abductive syllogism OOA-2 is possible to derive OOI-2 (those particular replicas are represented in italics). Consequently, from twenty-four abductive syllogisms, it is possible to reach thirty-two.

To provide examples of every one of these thirty-two abductive syllogisms would make this article too long. Then, we may limit it to some very precise examples. The syllogism AAA-2 can be instantiated with Peirce's example of beans and Schurz's example of law-abduction: if all the beans from this bag are white, and all these beans are white, therefore, all these beans are from this bag (EP1, 188). And if whatever contains sugar is sweet and all pineapple is sweet, therefore, all pineapple contains sugar [17]. The same premises can also suggest particular conclusions such as 'some of these beans are from this bag' and 'some pineapples are sweet', and so we have syllogisms AAI-2 as a replica of AAA-2.

Another example provided by Peirce is that which says: 'Only governors of Turkish provinces receive the honor of being surrounded by horsemen holding a canopy over his head'; 'In a Turkish province, there is a man upon a horseback, surrounded by four horsemen holding a canopy over his head'; and 'Therefore, that man is a governor of province' (see EP1, 189). This syllogism corresponds to AII-2, that is, a syllogism that is based on a fact (when Peirce landed in a seaport in Turkey saw such a man) and in a law (only governors are greatly honored); and it is a syllogism whose explanatory hypothesis is a fact (that particular man is a governor of province). It corresponds also to Thagard's simple abduction (1988, pp. 54–55) whose example is 'all rock-musicians dresses outrageously', 'Michael dresses outrageously', therefore, 'Michael is a rock musician.' It also corresponds to Thagard's existential abduction, which is, according to him the way in which the theory of phlogiston in chemistry was formulated: 'every object that contains an element that is given off, losses weight', and 'certain object loses weight', therefore, 'there should be such as element y' (see 1988, 57–58).

Other examples more strange are those with particular premises and universal conclusions, such as the abductive syllogism IIA-2. Thus if we know for certain that in previous cases rich women were murder by their philandering husband, we may conclude an explanatory law that says that "the murderer of every rich woman murdered should be her philandering husband". This kind of syllogism may seem an induction for it is generalizing particular cases in law, but remember that induction does not explain the facts, and in this syllogism the achieved law is explanatory. Perhaps the replica III-2 may seem more abductive, for it will reach a particular explanatory hypothesis. Thagard's examples of analogical abduction fit with these patterns. If a detective knows from a previous case that a rich woman was murdered by her philandering husband, then in a new case, if another rich woman was also found dead, he may suspect that this rich woman was also murdered by her philandering husband.

Other examples that differ greatly from deduction are those abductive syllogisms that have negative premises. For instance, we can imagine that if none crustaceous has a tracheal respiration system and none crab has a tracheal respiration system; then, perhaps crabs are crustaceous. This is an EEA-2 abductive syllogism. Or if the conclusion is particular such as 'this particular crab is a crustaceous, then we have an EEI-2 abductive syllogism.

Similarly, based on the information 'none planet bright by itself' and 'Mars does not bright by itself', it may be derived the conclusion 'Mars is a planet'. This is an EOI-2 abductive syllogism.

Likewise, it can be strange (from a deductivistic point of view) that from particular and negative premises something can be inferred. But abductively, this information can suggest some conclusion. For instance, if we know that 'Charles is not a friend of George', and 'John is not a friend of George either', then perhaps we may conclude that 'Charles and John are friends'. This is an OOI-2 abductive syllogism. And perhaps we may risk a more general conclusion and infer that 'those who are not friends of the same person are friends to each other', which correspond to an OOA-2 abductive syllogism.

These examples suffice to show that the internal conditions or criteria that allows to identifying types of abduction can be found in the syllogistic account of logical inferences. The procedure employed by Peirce at reversing the proposition of a deductive syllogism in order to explain abduction and induction can be extended to all moods and figure and find many more abductive syllogism up to a total of thirty-two.

Returning to the external condition of classification of types of abduction, let's recall that Peirce postulate time and consciousness as two criteria, and that, although he mentioned it, the guiding principles mentioned for further scholars can work as

Table 8 Four types of abduction according to two external conditions

	Urgent	Non-urgent
Conscious	1. Conscious and urgent abduction	2. Conscious and non-urgent abduction
Subconscious	3. Subconscious and urgent abduction	4. Subconscious and non-urgent abduction

a third external criterion. Time and consciousness are precise and understandable and were clearly explained above. I just say again that they are a continuous process without a clear line of demarcation. There is no marking point in which we can affirm clearly that a conscious act of thought became unconscious, and an urgent process became non-urgent. Holding this in mind, we can establish that these two distinctions are highly relevant to attain a classification of types of abduction. They are compatible with each other; thus a conscious hypothesis can be either urgent or not urgent. In consequence, from these two distinctions, four different types of abduction can be revealed (Table 8).

The third external criterion of classification depends upon the guiding principle that rules the inference. Peirce mentioned that every inference follows a guiding principle but he did not postulate different types of these principles. Instead, other scholars postulate the following principles: analogy [17, 19], known law, known scientific theory, extrapolation of background, speculation, causal unification [17], from a given database or selective; from a new personal explanatory hypothesis (P-creative); and from a new historical explanatory hypothesis (H-creative) [9].

All these criteria can be combined for they do not contradict with each other (Table 9).

We obtain thirty-two types of abductive inferences based on internal criteria and twenty-four types based on external criteria. These numbers can be mixed as well, for the internal distinctions, as we will see, are not contradictory with the external ones. In consequence, there can be seven hundred and sixty-eight (768) types of abduction.

To provide examples of every one of this copious variety of abductive inferences is a laborious task that would render this text extremely long. It will be a task of the community of inquirers to test whether all these abductive syllogisms are real or not. Hence, in order to conclude my inquiry I am going to provide just a few examples with the syllogism AAA-2. Above it was identified with the example of 'All pineapples contain sugar;' this particular example can be analyzed as a conscious and non-urgent inference, whose guiding principle is a law. This same form of syllogism AAA-2 (the same internal conditions) can also have different external conditions such as a subconscious and urgent inference guided by an analogy. For example, if someone observes (subconscious and uncontrolled perceptual judgment) that a warm and red liquid is flowing from an injury; and observes that that injured person lies on the floor unconscious can infer that that person was just been stabbed and she needs to be taken to a hospital right away.

Table 9 Twenty-four types of abduction from three external criteria of classification

	Conscious and urgent abduction	Conscious and non-urgent abduction	Subconscious and urgent abduction	Subconscious and non-urgent abduction
Analogy	Analogical conscious and urgent abduction	Analogical conscious and non-urgent abduction	Analogical subconscious and urgent abduction	Analogical subconscious and non-urgent abduction
Known laws	Conscious and urgent abduction from a law	Conscious and non-urgent abduction from a law	Subconscious and urgent abduction from a law	Subconscious and non-urgent abduction from a law
Known scientific theories	Conscious and urgent abduction from a scientific theory	Conscious and non-urgent abduction from a scientific theory	Subconscious and urgent abduction from a scientific theory	Subconscious and non-urgent abduction from a scientific theory
Extrapolation of background	Conscious and urgent abduction from an extrapolated background	Conscious and non-urgent abduction from an extrapolated background	Subconscious and urgent abduction from an extrapolated background	Subconscious and non-urgent abduction from an extrapolated background
Speculation	Speculative conscious and urgent abduction	Speculative conscious and non-urgent abduction	Speculative subconscious and urgent abduction	Speculative subconscious and non-urgent abduction
Causal unification	Conscious and urgent abduction from a causal unification	Conscious and non-urgent abduction from a causal unification	Subconscious and urgent abduction from a causal unification	Subconscious and non-urgent abduction from a causal unification

References

1. Aliseda, A.: Abductive Reasoning: Logical Investigations into Discovery and Explanation, p. 225. Springer, Dordrecht (2006)
2. Aristotle. A.: The basic works of aristotle. In: Mckeon, R. (ed.) The Modern Library, New York (2001)
3. Aristotle, A.: Categories, On Interpretation, Prior Analytics. Loeb Classical Library, no. 325. Translated by Harold P. Cooke and Hugh Tredennick. Harvard University Press, Cambridge, Massachusetts (1938)
4. Bonfantini, M., Proni, G.: To guess or not to guess?. In: Eco, Sebeok (eds.) The Sign of Three, pp. 164–184 (1988)
5. Campos, D.G.: On the distinction between Peirce's abduction and Lipton's Inference to the best explanation. Synthese **180**, 419–442 (2011)
6. Eco, U.: Horns, hooves, insteps: some hypotheses on three types of abduction. In: Eco, Sebeok (eds.) The Sign of Three, pp. 198–220 (1988)
7. Flórez, J.A.: Peirce's theory of the origin of abduction in Aristotle. Trans. Charles S. Peirce Soc. Q. J. Am. Philos. v.50 fasc.2, 265–280 (2014)

8. Flórez, J.A.: Peirce's reading on Aristotle's account of induction. In: Discusiones Filosóficas. Año 18 Nº 31, julio – diciembre, pp. 41–57 (2017)
9. Hoffmann, M.: Theoric transformations and a new classification of abductive inferences. Trans. Charles S. Peirce Soc. **46**(4), 570–590 (2010)
10. Magnani, L.: Abduction, Reason and Science: Processes of Discovery and Explanation. Springer, New York (2001)
11. Niño, D.: Abducting Abduction: Avatares sobre la comprensión de la abducción de Charles S. Peirce (Ph.D. Dissertation). Universidad Nacional, Bogotá, Colombia (2007)
12. Paavola, S.: On the Origin of Ideas: An Abductivistic Approach to Discovery (Ph.D. Dissertation.). University of Helsinki, Helsinki-Vantaa (Philosophical Studies from the University of Helsinki 15). https://www.researchgate.net/publication/47930944_On_the_Origin_of_Ideas_An_Abductivist_Approach_to_Discovery (2006)
13. Peirce, C.S.: The Charles S. Peirce Papers, Microfilmed Edition of C. S. Peirce's Manuscripts Preserved at the Houghton Library, Harvard University Library, Photographic Service, Cambridge, MA, 1966. Quoted following R. Robin, Annotated Catalogue of the Papers of Charles S. Peirce, University of Massachusetts Press, Amherst, MA (1967)
14. Peirce, C.S.: The Charles S. The Collected Papers of Charles Sanders Peirce. InteLex Corporation, Charlottesville, VA (1994)
15. Peirce, C.S.: The Charles S. The Essential Peirce: Selected Philosophical Writings. Nathan Houser and Christian Kloesel, vol. 2. Indiana University Press, Bloomington (1992)
16. Peirce, C.S.: The Charles S. Writings of Charles S. Peirce: A Chronological Edition, vol. 8. Indiana University Press, Bloomington, c1982<c2010>
17. Schurz, G.: Patterns of abduction. Synthese **164**(2), 201–234 (2008)
18. Thagard, P.: Semiosis and Hypothesis Inference in Ch. S. Peirce. Versus, pp. 19–20 (1978)
19. Thagard, P.: Computational Philosophy of Science. MIT Press, Cambridge, MA (1988)

Practical Abduction for Research on Human Practices: Enriching Rather Than Testing a Hypothesis

Sami Paavola

Abstract Following C. S. Peirce, abduction is often interpreted as a first phase of inquiry where a hypothesis is formulated requiring testing. I maintain, however, that a natural scientific ideal of testing is *not* the most suitable model for studies on human practices. Practical experimentation follows a different kind of a logic, and Peirce's formulations need to be developed further. I interpret abduction in relation to the Deweyan idea of a working hypothesis, and the method of ascending from the abstract to the concrete. Practical abduction is about enriching the working hypothesis instead of "testing" it in a strict sense. In studies on human practices abduction continues throughout the research process.

1 Introduction

Studies on human practices or activities have become widely adopted in various areas of research after 1990s [42]. There are different traditions within theories of practice with diverse methodological emphases [3, 27]. Richard Bernstein [1] distinguished four classic approaches to practice: (1) Hegelian tradition followed by Marx, (2) Peirce's and Dewey's pragmatism, (3) concept of practice in existentialism (like Sartre), and (4) in analytical philosophy. This distinction still provides a useful starting point for different philosophical approaches to practice theories which must be complemented with a plethora of theories influenced by sociological theories (like Bourdieu, Giddens), philosophy (especially Heidegger), and science and technology studies (especially actor-network theory) [23]. Despite a wide variety of approaches, it can be argued, that studies on human practices have some common characteristics. Most of the "practice-based research" or "practice theory" challenges basic dichotomies, or dualisms prevalent at discussions on methodologies, such as action versus structure, mind versus body, explanation versus understanding, qualitative versus quantitative research, determinism versus free will, subject versus object,

S. Paavola (✉)
Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland
e-mail: sami.paavola@helsinki.fi

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_2

deductive versus inductive approaches, or theory versus practice (Nicolini 2012, 2–3; [10], 3; [22]).

There are several classic methodological reconstructions or models on main phases (or aspects) of the research process such as the hypothetico-deductive model, inductivism, the method of analysis and synthesis, inquiry models, interrogative models, the hermeneutic circle, or experimentation. There is an old debate whether the methodological processes should lean either on natural sciences as giving a prototypical model for research, or focus on human sciences and specific features of human beings thus requiring completely different kinds of approaches compared to natural sciences. In his classic book "Explanation and Understanding", Finnish philosopher G. H. von Wright argued that there is a need for a simplified methodological model of its own for sciences of man which would give a "definitive alternative" to prominent models of natural sciences [46]. Von Wright maintained that natural sciences are "Galilean" emphasizing causality and the covering law model of explanation. Sciences of man, in turn, are continuing "Aristotelean" tradition with teleological explanations (ibid., 1–33). Natural sciences are targeted at "explanation" in the causal sense whereas human sciences at "understanding" and reasons behind action. According to von Wright human beings are intentional in their action following normative rules, and this requires a different methodological approach compared to natural sciences. Von Wright ended up suggesting that *practical syllogism* provides a methodological kernel for this kind of an alternative where the assent to the premises entails action (ibid., 26–27).

Following von Wright, it is argued in this paper that studies on human practices or activities are different to basic ideals of natural sciences emphasizing causal relationships, explanation, and the so-called covering law model, or subsumption theory of explanation (ibid., 1–33). A methodological reconstructions are needed that bring forth the special features of human practices, and ways of investigating them. But in contrast to von Wright an alternative to natural sciences in this paper is searched based on studies on transformation of practices rather than on understanding, and on "practical experimentation" rather than on practical syllogism and teleology. Practical experimentation and practical abduction are providing a better methodological reconstruction for practice research. The research on human practices can be seen as a third tradition [46], or a third "culture" (cf. [43]), and an alternative both to the ideal of natural sciences and to the ideal of hermeneutically interpreted human sciences.

The focus on practices or activities do not concern just the object of the research. It could also aim at taking better into account researchers' own methodological practices. Methodologies are always "reconstructions" rather than straightforward descriptions of the "logic-in-use", using classic formulations by Kaplan [17]. Practice-based approaches could help making these reconstructions more in line with real practices of research, and not just reconstructions along the ideals of research. Methodological models should not just represent a "logic of the finished research report" [13] but also help formulating the processes of research.

One aspect of methodology is to give paradigmatic or prototypical reconstructions of phases or aspects of research in order to give tools for planning, describing and justifying the research process. In this paper pragmatism and (cultural-historical) activity theory are used as a starting point for developing practice-based reconstruction of the phases or aspects of the research process. Both pragmatism and activity theory regard experimentation, transformative practical activity and intervention as essential parts of studying human practices [22]. Pragmatism and activity theory are natural candidates for providing a theoretical and philosophical basis for such practice-based methodology that is focusing on practical experimentation and transformations.

Abductive reasoning can provide one important methodological element for a methodological reconstruction of studying human practices. This requires, however, that abduction is interpreted and developed in relation to practice theories. Abduction was first developed by Charles S. Peirce in relation to his version of pragmatism. One standard interpretation is that abduction is the first phase of testing a hypothesis following basically the ideals of natural sciences ([31], 45–47). It can be maintained, however, that this is *not* the most suitable conception for practice research. "Practical experimenting" should be interpreted differently than the testing a hypothesis influenced by the natural sciences. In this paper, a form of abduction ("practical abduction") is developed further which would be more in line with the needs of practice research where theories and practices are enriching each other.

Peirce's own methodological ideals came from natural sciences and mathematics, and somewhat from historical research. Peirce's formulations, however, provide means to see the relevance of abduction also in the relation to studies on human practices (and practice research). Abduction related to practices has been used in empirical research influenced by cultural-historical activity theory, and pragmatism, for example, in relation to habits ([28], 77), to tools ([45], 40) to methods of an analysis "from above" and "from below" ([40], 61). A more philosophically oriented direction to develop abduction has been to interpret it within practical syllogism where the conclusion is the intention to do something, instead of a proposition [15, 45]. Magnani has suggested a specific form of abduction that he calls "manipulative abduction" which refers to "thinking through doing and not only, in a pragmatic sense, about doing" ([19], 229). Even if Dewey never referred to abduction as a specific form of an inference there lurks, however, a "Deweyan abduction" in several of his formulations of inquiry and reflective thought that may also contribute to the methodology of practice research [33].

In this paper it is maintained that the focus on practical experimentation has effects on the interpretation on abduction. In the first chapter, Peirce's basic notions of abduction in relation to his conception of pragmatism is presented. It is maintained that Peirce's formulations are more in line with the ideals of natural science than with practice research. In the second and third chapter the notions of *working hypothesis* and *practical experimenting*, as well as the *method of ascending from the abstract to the conrete* are analyzed as giving ingredients for practical abduction and practical experimentation. In the forth chapter the characteristics of practical abduction and differences to Peirce's typical formulations of abduction are analyzed.

2 Peirce's pragmatism and abduction

It is not at all obvious how Peirce interpreted abduction in relation to practices, or pragmatism. In his famous 1903 Harvard lectures Peirce made, however, a close connection between pragmatism and abduction.

> If you carefully consider the question of pragmatism you will see that it is nothing else than the question of the logic of abduction. That is, pragmatism proposes a certain maxim which, if sound, must render needless any further rule as to the admissibility of hypotheses to rank as hypotheses, that is to say, as explanations of phenomena held as hopeful suggestions; and furthermore, this is *all* that the maxim of pragmatism really pretends to do, at least so far as it is confined to logic, and is not understood as a proposition in psychology. (Peirce EP 2:234, 1903).

It is, however, well known that Peirce had conflicting views when it comes interpreting pragmatism as an approach emphasizing practical matters. This is most clearly seen in his famous 1898 lectures [37]. When asked to talk on "vitally important topics", Peirce was irritated and seemed to make a clear difference between scientifically and practically important matters: "… by my training I am nothing [but] a scientific man myself and am quite out of my element in talking about things vitally important" [37], n. 2 (deleted passage), 284). Dewey's pragmatism is often seen more promising when it comes to finding linkages between practice theories and pragmatism (see e.g. [20, 2]). It should also be remarked that Peirce's maxim of pragmatism is targeted at clarifying *concepts* (Peirce EP 1: 132, 1878) and not, for example, at clarifying or developing practices or developing concepts *in relation* to practices.

In his manuscripts Peirce formulated an inference called "Practical Retroduction" where retroduction is an alternative name for abduction (Peirce MS 637). The difference to "Scientific Retroduction" here was mainly, however, that in urgent practical situations the hypothesis must be chosen immediately: "A general who during a battle must instantly risk the existence of a nation either upon the truth of a certain hypothesis or else upon its falsity, must perforce go upon his judgment at the moment" (ibid.). Whereas in a scientific situation "indefinite time can be allowed for judging of the truth of an explanatory hypothesis" (ibid.). According to Peirce, in science, questions and hypotheses are always in principle open-ended and future-oriented, and they can be reconsidered. In practical situations the choices must be made more immediately.

One standard interpretation of abduction is that it is the *first phase* of inquiry where a hypothesis is formulated, followed by testing of the hypothesis with deductive and inductive means (EP 2:75–114, 1901; EP 2:440–442, 1908). Abduction is targeted at giving a plausible hypothesis but it does not give any security to its conclusions, and that is why the hypothesis must be tested (EP 2: 440). Testing means examination of the hypothesis, and deducing conditional experiential consequences which would follow from its truth (ibid., 441). The last stage is inductive "that of ascertaining how far those consequents accord with Experience, and of judging accordingly whether the hypothesis is sensibly correct, or requires some inessential modification, or must be entirely rejected (ibid., 442).

Peirce's emphasis on testing of a hypothesis is good as far it goes. It could be argued, however, that this kind of a model is simplifying also the picture of natural scientific testing. It gives a "clear-cut rationalist picture of experimentation as a theory-driven activity" ([41], 27) "missing multiple, epistemic, technical, institutional, and social resources involved in giving shape to experimental practice" (ibid., 140). But in any case, from the point of view of the study of practices, the ideal comes from a classic view on natural scientific experiments. There the truth of the hypothesis is assessed based on more or less clear causal relationships. One illustration of this model is the classic hypothetico-deductive model of inquiry [14].

This kind of testing may be relevant also in human (or social) sciences. But I would argue, that usually the function of concepts and theories is more instrumental in practice research from the start. Apparent examples are pedagogical models developed in educational sciences, or management theories in organizational sciences. They are not meant so much for giving strict hypotheses on definite expectations on the consequences to be tested but rather giving means also for practitioners for developing their practices.

3 Practical experimenting with the working hypothesis

Many formulations of practice research are actually quite close to the idea of hypothesis testing similar to the natural scientific ideal. One central idea within transformatively oriented practice-based research is that hypotheses need to be tested (or experimented) in practice (Dewey LW 8, 200–206; [24], 441–442; see also [7], 289). I argue, however, that the term testing is not the best analogy for practice research. It is important to see the differences of testing along the ideals of natural science (here called "natural scientific testing") with the ideals of "practical experimenting". The ideal of natural scientific testing aims at controlling and measuring the dependencies whereas (what is here called) practical experimenting is focusing on implementing and developing things in practice. In natural scientific testing the idea is to manipulate things in order to demonstrate the causal relationships whereas in the practical experimenting the idea is not basically to manipulate what is done but to observe how the practitioners themselves are modifying things [26].

Some formulations by Dewey seem to have a similar intepretation of the meaning of hypothesis testing than Peirce did. The testing means that on the basis of the hypothesis, it is observed if certain consequences follow or not. Dewey stated:

> The concluding phase is some kind of testing by overt action to give experimental corroboration, or verification, of the conjectural idea. Reasoning shows that if the idea be adopted, certain consequences follow. So far the conclusion is hypothetical or conditional. (Dewey LW 8.205; see also MW 6.236-7; MW 9.157; LW 8.200)

The different role of hypotheses starts to be depicted with the notion of a "working hypothesis". According to Dewey, the first suggestion is treated as a guiding idea, or a working hypothesis:

> The facts or data set the problem before us, and insight into the problem *corrects, modifies, expands* the suggestion that originally occurred. (Dewey LW 8, 202, 1933; the emphasis by SP)

Dewey's formulation above (*How We Think*, 1933) does not, however, draw a clear distinction to the natural scientific ideal. His examples in the passage above refer to a physician examining a patient or a mechanic inspecting a piece of complicated machinery (Dewey LW 8, 202). He presents the concluding phase of the reflective thought to be similar to the ideals of natural scientific testing:

> The concluding phase is some kind of testing by overt action to give experimental corroboration, or verification, of the conjectural idea. Reasoning shows that if the idea be adopted, certain consequences follow. (Dewey LW 8, 205)

The idea of the working hypothesis was introduced already 1899 by G. H. Mead in relation to social reforms [22]. According to Mead [21], hypotheses must be treated as "working hypotheses at the best" requiring further investigation, and being prepared to unexpected, because of complexity of forces influencing on social reforms.

> It is always the unexpected that happens, for we have to recognize, not only the immediate change that is to take place, but also the reaction back upon this of the whole world within which the change takes place, and no human foresight is equal to this.([21], 369–370)

In the paper Mead gives an impression that testing related to social reform is quite similar to testing by scientists in the laboratory. However, he also pointedly refers to one considerable difference to natural scientific testing. In social world we are potentially changing things which we at the same time assume to be somehow fixed:

> In the physical world we regard ourselves as standing in some degree outside the forces at work … In society we are the forces that are being investigated, and if we advance beyond the mere description of the phenomena of the social world to the attempt at reform, we seem to involve the possibility of changing what at the same time we assume to be necessarily fixed. (see ibid., 370–371)

In this short paper Mead does not analyze further the difference between testing in the laboratory and in relation to social reform but he pointed out that "reflective consciousness" by human beings makes a clear difference to natural sciences (ibid., 371).

Also Dewey had a formulation concerning experimentation where he is making a clear difference between interpretations concerning (natural scientific) laboratories and "method of social matters":

> When we say that thinking and beliefs should be experimental, not absolutistic, we have then in mind a certain logic of method, *not, primarily, the carrying on of experimentation like that of laboratories* [emphasis by SP]. Such a logic involves the following factors: First, that those concepts, general principles, theories and dialectical developments which are indispensable to any systematic knowledge be shaped and tested as tools of inquiry.

> Secondly, that policies and proposals for social action be treated as working hypotheses, not as programs to be rigidly adhered to and executed. They will be experimental in the sense that they will be entertained subject to constant and well-equipped observation of the consequences they entail when acted upon, and subject to ready and flexible revision in the light of observed consequences. (Dewey LW 2.361–362)

This formulation is more in line with requirements for practical experimentation than the idea of the natural scientific model of testing a hypothesis. According to Dewey, "concepts, general principles, theories and dialectical developments" are seen as *tools* of inquiry (not hypotheses to be accepted or rejected) and they are treated as working hypotheses instead of "programs to be rigidly adhered to and executed". The emphasis on *working hypotheses* means that the hypothesis is tentative but also that the hypothesis itself is worked with during the process.[1] Dewey is talking also here on "observed consequences" but an important addition is that proposals or hypotheses are "subject to ready and flexible revision" on the basis of constant observation of the consequences.

What is then a difference between natural scientific testing and practical experimenting? I think that there are at least two main differences in Dewey's as well as Mead's alternative formulations on working hypotheses. The first one is about the constant readiness to modify the working hypothesis during the research. In this sense, the abductive "phase" typically continues throughout the research process. Practical experimenting requires that the lessons learned from the practical experimentation should be taken into account in the conceptual apparatus. It is not so much about testing if a certain specific kind of a conceptual idea works (or not) but about developing the conceptual idea during the experimentation that seems to capture essential elements of the phenomena or of the changes studied. And as Mead pointed out, it is also unexpected reactions from the world which are potentially transforming the original idea during the experimentation. The second difference is about the role of a conceptual idea in a comparison to activities and observations. In the ideal of the natural scientific testing, the idea is to deduce consequences from the hypothesis. Practical experimentation is focusing on "transformational practical activity", being close to an interventionist approach [22]. Here experimentation requires not only new practically oriented concepts and organizational models but also new instruments and tools [24], and agency by practitioners [8] oriented at transforming the material interaction with the environment. The role of tools and cultural remediation is often neglected by pragmatism but it is emphasized in the cultural-historical activity theory ([25], 118–119; [24]). Practical experimentation is then not just about testing conceptual ideas. Rather it is about giving means for developing tools and instruments with related agency, and inducing the change in practice along with conceptual development. This kind of research is not just about implementing a certain kind of an idea but besides that and throughout the process developing novel conceptual and practical means that would support the development (See also [7]).

[1]Reijo Miettinen (a personal conversation) suggested this meaning for the term "*working* hypothesis". An implementation of it requires developmental work by practitioners themselves.

4 From testing to enriching: the method of ascending from the abstract to the concrete

In the tradition of cultural-historical activity theory there is another way of describing what is happening in this kind of research besides using the notions of testing or experimentation. The "method of ascending from the abstract to the concrete" is based on Marx and dialectics ([13, 3]; [4], 243–289). It has *not* been much applied in empirical research but rather provided one basis, and an important exemplar, for activity theoretical formulations of methodology (see [25], 112–117; [7]). For practice-based research this method gives fertile means of developing methodological formulations further besides its dialectical uses [44].

It is not easy to present the method of ascending from the abstract to the concrete briefly because both Ilyenkov's and Davydov's presentations of it are conceptually intricate descriptions with many methodological sub-themes. Miettinen (2000, 111–112) has analyzed this method in relation to activity theory and basic formulations by Marx, and presented it succinctly with three basic steps:

The study starts from the concrete "chaotic whole".

It descends, then, to the abstraction of the basic determining categories.

Thirdly, it "rises" again – using the abstraction – to the concrete whole, this time as "a rich totality of determinations and relations"

Dafermos ([4], 258) depicts the same process dialectically as "a contradictory unity of two distinct and opposite movements of thinking: a) the movement from sensory-concrete perception to abstract thinking and b) the ascent from abstract thinking to the mentally concrete and from mentally concrete to practice."

Marx' analysis has been both politically controversial and heated but also giving a very demanding model for an empirical research (his own analysis of capitalism took decades) (ibid., 112). Still, it gives one fertile starting point for understanding the relationship between theoretical concepts and practice. The movement from sensory concrete to abstract thinking aims at constructing abstract concepts or hypotheses to be developed further. This is similar to abduction, that is, to produce a promising hypothesis related to the "chaotic whole". The further aim is not just to test these concepts or hypotheses but to refine and enrich the abstraction in relation to practices (with "mentally concrete").

An important starting point for Davydov is a difference between an empirical interpretation of a generalization, and a "theoretical generalization" ([5], 130–137; see also [39], 84–92). The empirical generalization is based on a "transition from a description of the properties of a particular object to finding and singling them out in a whole class of similar objects" (ibid., 5, 18). This does not capture a theoretical generalization which is based on a different kind of an abstraction. The theoretical generalization is based on an analysis of an individual, concrete case: "In reality it always happens that a phenomenon which later becomes universal originally emerges as an individual, particular, specific phenomenon, as an exception from the rule" ([16], 83). A theoretical generalization aims at finding essential or real features and relationships of things. It reproduces developmental processes and also possibilities of certain things and their relationships. It finds connections between things which before having that generalization are fragmentary or undeveloped.

The distinction between an empirical generalization and a theoretical generalization has similarity, or at least a close affinity with Peirce's early formulations of a distinction between induction and abduction (respectively). In his early papers Peirce made a difference between induction and abduction with syllogistic formulations. Induction is the inference of the major premise (rule) from the minor premise (case) and the conclusion (result), and abduction the inference of the minor premise (case) from the conclusion (result) and the major premise (rule) (Peirce CP 5.275–276, 1868; CP 2.623, 1878). In his later works Peirce broadened his conception of induction but never abandoned his early syllogistic formulations. Typical induction, in line with this syllogistic formulation, is then about making an empirical generalization.

Case.—These beans are from this bag.
Result.—These beans are white.
Rule.—All the beans from this bag are white.

Abduction is different. It starts (usually) with noticing some curious or anomalous phenomena, and then finding a hypothesis that would explain them, or make understandable things which at first seemed detached with each other (Peirce EP 1: 188–189, 1878; EP 2: 231, 1903). As N. R. Hanson stated it later, abduction is about finding a novel, intelligible or organized "pattern" to solve puzzling phenomena [11, 12]. Induction is based on quite straighforward empirical generalization whereas abduction requires skills of a detective, that is, by analyzing details and specificities of a case, it aims at finding a qualitatively novel solution that makes anomalous things understandable and gives means to anticipate some other things.

Abduction and the whole area of the logic of discovery was for long neglected in the philosophy of science and in methodological literature [31]. There is a similar tendency in the discussion on the method of ascending from the abstract to the concrete to neglect the phases of discovery even by Davydov ([5], 129):

> With what must such reproduction begin? The very name of its method indicates that one must go from the abstract, and, actually, "abstract definitions lead to reproduction of the concrete by means of thought"

One reason for this neglect might be that these processes of discovery are often reconstructed afterwards (see [7], 304).

Usually Davydov, however, is emphasizing the whole process, that is, both the analysis and the way of delineating the abstraction, and the *ascent* from the abstract to the concrete again ([5], 137):

> In other words, the formation of a conceptual generalization presupposes not only a transition from the concrete and individual to the abstract and general, but also the reverse transition from the general and the abstract to the individual and the concrete. (ibid., 12)

In this sense, the whole name of the method should be something like "movement from the sensory concrete to the abstract, and ascending from the abstract to concrete in practice". This would, however, be clumsy as a name!

Yrjö Engeström has interpreted the method of ascending from the abstract to the concrete for developing the cycle of expansive learning in developmental work research ([7], 288–289; see also [25], 114–117). In the theory of expansive learning the first phases of collective concept formation includes questioning ("questioning, criticizing or rejecting some aspects of the accepted practice") and analyzing the situation (both tracing the origination and evolution of the activity in question, and analyzing the actual-empirical situation) which are oriented at developing a model of a new idea (a "germ cell") ([7], 288–289). In the cultural-historical activity theory the phases of producing the germ cell include detailed analyses of the activity in question from different perspectives and timescales (see also [25], 114–117; [4], 256–259). This is in line with the pragmatist emphasis on defining present activities fundamentally in relation to the past and to the future, that is, with the irrevocable past and heading for the future, even though in the pragmatist tradition the historical analysis is not explicitly included in the method as is the case in the cultural-historical activity theory ([25], 118).

The phase of ascending from the abstract to the concrete along the model of expansive learning differs from the Peircean model of testing a hypothesis. It is not a question of an individual hypothesis that is tested (in a Peircean sense) but practical experimenting means that a broader model is developed which is then implemented in practice and reflected and developed further. The "germ cell", a specific kind of a working hypothesis, has the essential features or relationships of the new model or a solution that needs to be practically implemented and developed in the process of a synthesis ([7], 293). It means that the germ cell idea is transformed into practice; which is not a simple process.

> A theoretical concept is initially produced in the form of an abstract, simple explanatory relationship, a "germ cell." This initial abstraction is step-by-step enriched and transformed into a concrete system of multiple, constantly developing and expanding manifestations. In other words, the initial simple idea is transformed into a complex new form of practice. ([7], 288)

In this sense, the germ cell is enriched and it "opens up a perspective for multiple applications, extensions and future developments (ibid., 289). The practical application means that the working hypothesis must be complemented with relevant tools and procedures to be tested in practice [25]. Engeström et al. [7] highlight in their research case of the home care workers that this "moving to the concrete" means embodiment and physical enactment. It is then not just about testing an idea or working with verbal definitions but about putting the idea in practice in concrete and tangible ways. In this case these enactments consisted on physically performing and/or examining movements or exercises needed by the new concept, developing artifacts of the environment to enhance the concept formation, and using bodily gestures and facial expressions to enhance an observation, idea, or feeling (ibid. 302–303). Miettinen ([25], 123, 126) has emphasized that the initial hypothesis is first an abstraction which needs to be transformed to a "testable working hypothesis", or "practical germ cell" to make it practically experimental.

5 Characteristics of practical abduction

How "practical experimenting" is then having an effect on the concept of abduction? Or is abduction in itself the same, concerning the way how hypotheses are formulated, but embedded to research on practices and transformation of practices? I maintain that practical experimenting has an effect at least on the overall methodological process, and also on the way how abduction is involved in that process. But how to interpret then practical abduction?

It must be remarked that there are several ways of interpreting abduction as developed by Peirce and then after Peirce (see [31], 21–55). Besides syllogistic interpretation (see above) and as a first phase of inquiry, Peirce treated abduction to be the same, or close to: (1) guessing, (2) insight, (3) perception and perceptual judgment, (4) sensations and emotions, (5) conceptions, (6) pattern recognition, (7) the maxim of pragmatism, (8) the economy of research, (9) interrogation, (10) inference of a cause from its effect, (11) the category of Firstness, (12) an inference through an icon, (13) pure play, and musement (see ibid., 46–47). These interpretations do not necessarily exclude each other but they show different directions for interpreting and developing abduction.

Peirce had some formulations related to abduction which have affinities to the method of ascending from the abstract to the concrete, like a formulation of the maxim of pragmatism:

> The elements of every concept enter into logical thought at the gate of perception and make their exit at the gate of purposive action; and whatever cannot show its passports at both those two gates is to be arrested as unauthorized by reason. (EP 2:241, 1903)

This is in itself a simple formula but there is an analogy to a methodological interpretation of the method of ascending from the abstract to the concrete (see Fig. 1) Both of them are emphasizing practices, first, by focusing on "purposive action" (or "transformative" practices", and secondly, by interpreting concepts operating in between perception and purposive action.

My own main interpretation of abduction highlights the dynamic nature of it. Abduction is a "weak" mode of inference (basically it allows inferences only on what "may be") but if used skillfully it has high "uberty" (Peirce's own term for "fruitfulness" or "productiveness") (Peirce CP 8.384–388, 1913; Peirce EP 2:472, 1913). The fruitfulness of abduction is not clear only by looking at "definitory rules" of abduction (i.e. what kinds of inferences are permitted with it) but by looking at "strategic rules" of using abductive reasoning [29]. In games like chess definitory rules are the basic rules of the game (for example, how pieces can be moved) but just by knowing these rules no one is good at playing chess. That requires skills of playing well, that is, strategic rules. At abductive reasoning, these strategies highlight the process of putting together abductive arguments while searching fertile or promising hypotheses. Abduction is reasoning targeted at search processes. These kinds of abductive strategies include, for example: search for somehow anomalous phenomena, observing details and little clues, continuous search for hypotheses

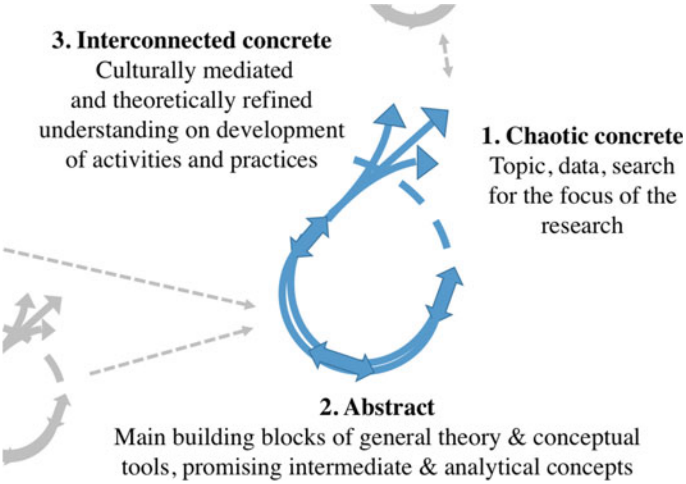


Fig. 1 Simplified elements of a methodological cycle, starting from "chaotic concrete" and searching analytical concepts to help anticipate and enrich transformative activities and practices. *Adapted from* Vetoshkina and Paavola [44]

(or novel explanations, or interpretations), delineating what *kind* of a hypothesis is searched for, aiming at finding explanations that can themselves be explained, search for "patterns" and connections that fit together ([31], 206–211). These strategies give elements for a "logic of discovery" opening up those processes that both constrain and instigate the search for novel ideas [31]. This kind of a "Hansonian" interpretation of abduction highlighting logic of discovery and a play with possibilities is different to interpretations of abduction giving *justification* to explanatory hypotheses [30].

A strategic approach to abduction highlights its holistic nature. Hypotheses are not typically searched for only in relation to just individual observations or anomalies but by taking into account different kinds of clues and constraints, both theoretical and empirical. This fits with studies on practices which are typically aimed at understanding complex changes of activities happening in different time-scales and in different "layers" or elements of activities (for example, situational activities, changes in organizations, changes in national or international rules and guidelines) (see [44]). In studies on practices working hypotheses are aimed at taking into account these different elements, or constraints from different directions which makes the formulation of the working hypotheses also qualitatively challenging.

Peirce, especially in his later writings, emphasized that human beings must have some kind of a guessing "instinct", or an insight for finding fertile hypotheses (Peirce CP 5.173, 1903). Peirce was speculating that when human mind is developed under the influence of laws of universe

> it is to be expected that he should have a *natural light*, or *light of nature*, or *instinctive insight*, or genius, tending to make him guess those laws aright, or nearly aright. (Peirce CP 5.604, 1903)

He then speculated on some kind of an "il lume naturale" (a sort of natural bent) (Peirce CP 1.80–81, c. 1896; CP 1.630, 1898).

Peirce more or less neglected cultural issues, and issues that are nowadays emphasized by distributed cognition as a basis for abduction (see [32]). This would be a natural extension to practical abduction. A culturally oriented alternative is to develop abduction based on "il lume culturale" rather than on il lume naturale ([2], 134; [31], 196). From the point of view of practice research abductive search is not done just by insightful individuals but collaboratively, using various kinds of cultural means, and in a historically changing, material context giving clues and ingredients for new ideas [32].

In studies of human practices conceptualizations are not separate from practices. Conceptualizations are engendered in close relationship to practices, and while implementing novel solutions in practice. They are not so much targeted on testing a conceptual idea but developing conceptualizations all along the research process in relation to practices. I think this is one focal meaning of the concept "working hypothesis" (or the "germ cell"). It means that also the hypothesis itself is under construction, and potentially re-evaluated. A working hypothesis is giving a rough outline for a solution that is supposed to be specified and developed during the process. This fits nicely with a strategic emphasis on abduction but *not* with the idea that abduction is just the "first" phase of inquiry. In this kind of practical abduction, abductive search for better conceptual means of understanding the situation in question continues throughout the research process. Often it is as important to figure out key problems and issues in the situation as finding solutions. This means then that the abductive process continues after the "first" phase of finding a promising hypothesis.

Studies of human practices emphasize reflection and evaluation throughout the research process. It is not just researchers who are supposed to be reflecting but also practitioners involved. This also means that

> concepts do not precede action nor do they legitimate actions only after the fact. They serve as means of making visible already developed possibilities, contradictions and emergent alternative solutions of a practice ([24], 442)

Engeström has emphasized (by citing Leontjev) that goals of action are often known only after action and as a result of a long process of acting ([7], 291)[2]. Similarly the development of conceptualizations is a long process done in a close interaction with experiments and actions developed in practice.

Peirce did not have the concept of a "working hypothesis" in his writings, at least as far as I know. But there is an affinity to N. R. Hanson's treatment of Peircean abduction as the logic of discovery. According to Hanson, an important element of the process of discovery is that the *type* of the solution can be delineated before the solution is acquired [13]. The aim is to find an organized *pattern* which takes into account all relevant issues involved [12]. This process of delineating the type of the hypothesis has clear similarities to functions of the working hypothesis. It gives an

[2]Leontjev ([18], 101) is here citing Hegel saying that the individual "cannot determine the goal of his acting as long as he has not acted …"

outline for a solution that is then targeted to be specified. The working hypothesis can be modified quite a lot during the process if the novel idea forms a relevant new totality or "pattern".

As I see it, practical abduction does *not* need to be tied to explanatory issues. Peirce had several basic descriptions of abduction where it was closely related to seeking or adopting *explanations*:

> Hypothesis [abduction] is where we find some very curious circumstance, which would be *explained* [my emphasis] by the supposition that it was a case of a certain general rule, and thereupon adopt that supposition (Peirce EP 1: 189, 1878).
>
> … the operation of adopting an *explanatory* [my emphasis] hypothesis, - which is just what abduction is … (Peirce EP 2: 231, 1903)

In his early writing Peirce also analyzed abduction as an inference of a cause from its effects (Peirce W 1:180, 1865). But all basic formulations of abduction are not tied to explanatory issues, or to seeking causes. Even in the well known formulation itself, explanation is not mentioned (even if it was present in his description of this formulation) (See Peirce EP 2: 231, 1903):

> The surprising fact, *C*, is observed;
>
> But if *A* were true, *C* would be a matter of course,
>
> Hence, there is reason to suspect that *A* is true.

This is in line with Gabbay and Woods [9] who have emphasized that abduction should not cover *just* explanatory accounts but more broadly transformations of "ignorance problems".

This kind of an non-explanatory account of abduction (see [20]) is important for practice-based research. I would argue that in the study of human practices the search for hypotheses and novel conceptualizations is not just about seeking explanations or causes. It is more about conceptualizing key issues in the phenomena in question, or being able to conceptualize key transformations in activities in question, which are targeted at developing these phenomena further. These working hypotheses or germ cells can help making explanations or even point out causes but the purpose of them is not just that. The idea of natural scientific testing is based on the idea that some kind of a causal relationship or an entity is presumed to be involved. But in the practical experimenting the main idea is rather if some kind of a working hypothesis is promising enough to be carried out and developed further.

One difference between basic formulations of abduction and practical abduction is the source of research problems, or puzzles to be studied. The basic formulations of abduction are bringing forth some kind of an anomaly in the existing knowledge that instigates the inquiry. The focus is on an intellectual process. The practice-based research emphasizes the central role of practitioners in the process of formulating the research problems. In the model of expansive learning the germ cell is based on analyzing contradictions within and between different activities and activity systems [7]. In Deweyan pragmatism and in activity theory the commitment to the problems and perspectives of the people and activities studied is a constitutive and also an ethical feature of the research ([22], 401–402). It means that people involved in the research have a bigger role at defining the research questions and problems than in the basic formulations of abduction. But there is no strict dichotomy here. Also in the practice-based research, the researcher has (or researchers have) a central role in formulating research questions and research focuses but it is typically done in close interaction with practitioners.

6 Conclusion

There is a need to develop novel methodological frameworks or reconstructions taking into account specifics of what is called here "studies of human practices", especially focusing on transformative practical activity. At the introduction of this paper I referred to G. H. von Wright's distinction between "explanation" and "understanding" providing, according to him, main models for natural sciences and sciences of man respectively. Von Wright himself remarked, in passing, that Hegel's and Marx's approaches are providing an anomaly to this distinction ([46], 7–8). They seem to have elements that are both close to the ideals of natural sciences emphasizing laws and causality, and to ideals of human sicences with teleological and intentional explanations. The overall argument of this paper is to maintain that practice-based research does not fit to the dichotomy between explanation and understanding and requires an area of its own on "practices". Here the aim of the research is not "explanation" in line with the covering law model, nor mainly hermeneutic "understanding" but rather finding grounds, means, or anticipations for transformative practices. It is based on transformative ontology (cf. [22]) where practices both resist and enable human activities.

The Peircean abduction provides fertile additional means for understanding studies of human practices but not by using standard models of abduction but by developing formulations of abduction further. Peirce himself was developing abduction in relation to a traditional picture of a natural scientific model of testing a hypothesis. It works with cases where there are clear-cut causal relationships involved or tested. The language of "testing" is used also in practice-based research, and with good reasons (e.g. Dewey LW 8, 205; [24], 442; [26]). It would be, however, important to see a difference between "practical experimenting" within sciences of man, to the ideals of "natural scientific testing".

I have searched elements for practical experimenting from basic procedures of the method of ascending from the abstract to the concrete used in cultural historical activity theory, and by developing further Dewey's (and Mead's) idea of a working hypothesis. Working hypotheses have several meanings in this kind of a reconstruction: 1. They are tentative (they can be accepted or rejected on the basis of the evidence). 2. They are themselves something that can be modified during the research process (some kind of drafts that are meant to be refined or modified on the basis of the practical experimentation). 3. They suggest activities which require work and targeted efforts by those who are using them by developing tools, ways of working, and related agency. The model of natural scientific testing is emphasizing mainly the first one of these but practical experimenting requires the other two besides.

The focus on working hypotheses in relation to practical experimenting brings together the function of practitioners in this kind of a research (what they are doing) on the one hand, and the epistemological questions concerning the research on the other hand. The working hypotheses are about transforming existing practices (by the practitioners) but on the other hand they are tools for researchers to interpret these practices from the point of view of the research.

Practical experimenting means that concepts are used more instrumentally than in the ideals of natural scientific testing. This is not a comment on the truth-value which these concepts are providing or not providing, but rather of the function of the concepts as a part of human activities. Concepts are supposed to give practical and theoretical means and tools for subsequent actions (for practitioners and for research) not just stating hypothesis on the state of affairs. This also changes the function of abduction in the research process. Maybe a bit suprisingly, abductive elements are more prevalent in studies on human practices than in the ideal of natural scientific testing. This is because the search for better conceptualizations continues often throughout the research process. In the natural scientific testing hypotheses are meant to be tested (with observations) but in practice research they are more like working hypothesis to be developed further in relation to emerging practices.

As I see it, the function of methodological reconstructions like practical abduction or practical experimenting is to give means for reconstructing and justifying also those fuzzy processes of research where conceptualizations are engendered and remoulded, not just help formulating the "logic of finished research report" (cf. [13]). These kinds of methodological models are still, however, reconstructions. There should be more empirical research on methodological processes of research in social and educational sciences (see [44]). I maintain that this kind of a model of practical abduction would in any case take better into account the characteristics of practise-based research than standard formulations of abduction.[3]

References

1. Bernstein, R.: Praxis and action. Contemporary philosophies of human activity. University of Pennsylvania Press, Philadelphia (1971)
2. Bonfantini, M.A., Proni G.: To guess or not to guess? In: Eco, U., Thomas, A.S. (eds.) The Sign of Three. Dupin, Holmes, Peirce, pp. 119–134. Indiana University Press, Bloomington (1983)
3. Buch, A., Schatzki, T. (eds.): Questions of Practice in Philosophy and Social Theory. Routledge, New York (2018)
4. Dafermos, M.: Rethinking cultural-historical theory. A dialectical perspective to Vygotsky (Vol. 4). Springer (2018)

[3]I want to thank Liubov Vetoshkina, Jaakko Virkkunen, and especially Reijo Miettinen for giving many invaluable comments and suggestions to this paper that definitely made the article better. All misinterpretations and omissions are my own.

5. Davydov, V.V.: Soviet Studies in Mathematics Education, vol. 2. Types of Generalisation in Instruction. Published by the National Council of Teachers of Mathematics, Reston Virginia (1990)
6. Dewey, J.: The collected works of John Dewey, 1882–1953. In: Boydston, J.A. (ed.) Southern Illinois UP. Published as The Early Works: 1991–1898 (EW), The Middle Works, 1899–1924 (MW), and The Later Works, 1925–1953 (LW). Citations are made with these designations followed by volume and page number (1969–1991)
7. Engeström, Y., Nummijoki, J., Sannino, A.: Embodied germ cell at work: building an expansive concept of physical mobility in home care. Mind Cult. Act. **19**(3), 287–309 (2012)
8. Engeström, Y., Sannino, A.: Studies of expansive learning: foundations, findings and future challenges. Edu. Res. Rev. **5**(1), 1–24 (2010)
9. Gabbay, D.M., Woods, J.: The Reach of Abduction Insight and Trial. A Practical Logic of Cognitive Systems, vol. 2. Elsevier, Amsterdam (2005)
10. Gherardi, S.: How to Conduct a Practice-Based Study. Problems and Methods. Edward Elgar, Cheltenham, UK (2012)
11. Hanson, N.R.: Patterns of Discovery. An Inquiry into the Conceptual Foundations of Science. University Press, Cambridge (1958)
12. Hanson, N.R.: Notes toward a logic of discovery. In: Bernstein, R.J. (ed.) Perspectives on Peirce, pp. 42–65. Yale University Press, New Haven and London (1965)
13. Hanson, N.R.: Is there a logic of scientific discovery?. In: Feigl, H., Maxwell, G. (eds.) Current Issues in the Philosophy of Science, pp. 20–35. Holt, Rinehart and Winston, New York (1961)
14. Hempel, C.G.: Philosophy of natural science. Prentice-Hall, inc., Englewood Cliffs, N.J., London (1966)
15. Hilpinen, R.: On practical abduction. Theoria **73**(3), 207–220 (2007)
16. Ilyenkov, E.V.: The Dialectics of the Abstract and the Concrete in Marx's Capital. First published 1960. Translated by Syrovatkin, S. Aakar Books Classics (2017/1960)
17. Kaplan, A.: The Conduct of Inquiry. Methodology for Behavioral Science. Chandler Publishing Company, San Francisco (1964)
18. Leontjev, A.N.: Activity, consciousness and personality. Englewood Cliffs, Prentice Hall (1978)
19. Magnani, L.: Model-based and manipulative abduction in science. Found. Sci. **9**(3), 219–247 (2004)
20. Magnani, L.: Abductive Cognition. The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning. Springer-Verlag, Berlin Heidelberg (2009)
21. Mead, G.H.: The working hypothesis in social reform. Am. J. Sociol. **5**(3), 367–371 (1899)
22. Miettinen, R.: Epistemology of transformative material activity: John Dewey's pragmatism and cultural-historical activity theory. J. Theor. Soc. Behav. **36**(4), 389–408 (2006)
23. Miettinen, R., Paavola, S., Pohjola, P.: From habituality to change: contribution of activity theory and pragmatism to practice theories. J. Theor. Soc. Behav. **42**(3), 345–360 (2012)
24. Miettinen, R., Virkkunen, J.: Epistemic objects, artefacts and organizational change. Organization **12**(3), 437–456 (2005)
25. Miettinen, R.: Ascending from the abstract to the concrete and constructing a working hypothesis for new practices. In: Oittinen, V. (ed.) Evald Ilyenkov's Philosophy Revisited, pp. 111–129. Series B: 13. Kikimora Publications, Helsinki (2000)
26. Muurinen, H.: Pragmatismi ja kokeileva lähestymistapa sosiaalityön tiedonmuodostuksessa. Doctoral Dissertaion, In Finnish. Heikki Waris-instituutin tutkimuksia 2/2019. Helsinki. Valtiotieteellinen tiedekunta (2019)
27. Nicolini, D.: Practice theory, work, and organization: An introduction. OUP Oxford (2012)
28. Norros, L.: Acting Under Uncertainty. The Core-Task Analysis in Ecological Study of Work. VTT Publications 546, Espoo (2004)
29. Paavola, S.: Abduction as a logic of discovery: the importance of strategies. Found. Sci. **9**(3), 267–283 (2004)
30. Paavola, S.: Hansonian and Harmanian abduction as models of discovery. Int. Stud. Philos. Sci. **20**(1), 91–106 (2006)

31. Paavola, S.: On the Origin of Ideas. An Abductivist Approach to Discovery, Revised and Enlarged edn. Lap Lambert Academic Publishing, Saarbrücken (2012)
32. Paavola, Sami, Hakkarainen, Kai: Three abductive solutions to the Meno Paradox—with instinct, inference, and distributed cognition. Stud. Philos. Educ. **24**(3–4), 235–253 (2005)
33. Paavola, S.:. Deweyan approaches to abduction? In: Zackariasson, U. (ed.) Action, Belief and Inquiry. Pragmatist Perspectives on Science, Society and Religion, pp. 230–249. Nordic Studies in Pragmatism 3. Nordic Pragmatism Network, Helsinki. Available http://www.nordprag.org/nsp/3/Paavola.pdf (2015)
34. Peirce, C.P.: Collected Papers of Charles Sanders Peirce, vols. 1–6, Hartshorne, C., Weiss, P. (eds.), vols. 7–8, Burks, A.W., (ed.). Harvard University Press, Cambridge, Mass (1931–1958)
35. Peirce, E.P. (1998) The Essential Peirce: Selected Philosophical Writings, vol. 2 (1893–1913), The Peirce Edition Project (eds). Indiana University Press, Bloomington and Indianapolis
36. Peirce, M.S.: Manuscripts in the Houghton Library of Harvard University, as identified by Richard Robin, Annotated Catalogue of the Papers of Charles S. Peirce. University of Massachusetts Press, Amherst (1967)
37. Peirce, C.S.: Reasoning and the Logic of Things. Ketner, K.L. (ed.). Harvard University Press, Cambridge (1992)
38. Peirce, W.: Writings of Charles S. Peirce: A Chronological Edition, 6 vols. (to date), The Peirce Edition Project (eds.). Indiana University Press, Bloomington (1982–)
39. Pihlaja, J.: Learning in and for Production. An Activity-Theoretical Study of the Historical Development of Distributed Systems of Generalizing. Yliopistopaino, Helsinki (2005)
40. Rainio, A.P.: Lionhearts of the Playworld: An Ethnographic Case Study of the Development of Agency in Play Pedagogy. University of Helsinki, Institute of Behavioural Sciences, Studies in Educational Sciences 233, Helsinki (2010)
41. Rheinberger, H.J.: Toward a history of epistemic things. Synthesizing proteins in the test tube. Stanford University Press, Stanford, CA (1997)
42. Schatzki, T.R., Knorr Cetina, K., von Savigny, E. (eds.): The Practice Turn in Contemporary Theory. Routledge, London and NY (2001)
43. Snow, C. P.: The two cultures and the scientific revolution. Cambridge University Press (1959)
44. Vetoshkina, L., Paavola, S.: From the abstract to the concrete and beyond: the winding road of constructing a conceptual framework. A manuscript submitted to a special issue in *Outlines* (2020)
45. Virkkunen, J.: Developmental interventions in work activities—an activity theoretical interpretation. In: Kontinen, T. (ed.) Development Intervention: Actor and Activity Perspectives. Center for Activity Theory and Developmental Work Research, Helsinki (2004)
46. Von Wright, G.H.: Explanation and Understanding. Cornell University Press, Ithaca (1971)

Abductive Spaces: Modeling Concept Framework Revision with Category Theory

Rocco Gangle, Gianluca Caterina, and Fernando Tohmé

Space is the place.
Sun Ra, cosmic jazz maestro

Abstract A formal model of abductive inference is provided in which abduction is conceived as expansive and contractive movements through a topological space of theoretical and practical commitments. A pair of presheaves over the (Heyting algebra) space of commitments corresponds to communities sharing commitments on the one hand and possible obstructions to commitments on the other. In this framework, abductive inference is modeled by the dynamics of redistributed communities of commitment made in response to obstructive encounters. This semantic-pragmatic model shows how elementary category theory tools can be used to formalize abductive inference while hewing close to ordinary intuitions about collective agency and reasoning.

1 Introduction

When avant-garde jazz pianist and orchestra director Sun Ra declares "space is the place" on the album for the 1974 film by the same name, he expresses a thought that is close to being formally tautological and yet offers significant theoretical insight.

R. Gangle (✉) · G. Caterina
Endicott College, Beverly, MA 01915, USA
e-mail: rgangle@endicott.edu

G. Caterina
e-mail: gcaterin@endicott.edu

F. Tohmé
UNS-CONICET, Bahia Blanca, Argentina
e-mail: ftohme@gmail.com

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_3

On the one hand, the concept of "place" or "locus" is, more or less by definition, relative to the concept of some "space" that envelops and determines it. "Places" in the plural are possible only with respect to the unity of some "space" within which they find their places with respect to one another. On the other hand, this banal or obvious coordination of places within spaces (and, in particular, the overarching unity and singularity of natural or cosmic space as such) serves as one of the most fundamental aspects of human experience and thereby becomes a nearly endless source for metaphors and figurative modes of orientation within thought. We encourage students to think "outside the box"; we deliberate "narrowly" or "broadly" about some given topic; a philosopher is said to "move" from one doctrinal "position" to another; a scientist considers a problem from a "higher point of view" or "flattens out" certain difficulties, and so forth.

It is ultimately the structural integration of local places within or according to the order of more general spaces that provides the basis for the use of "space" as a general source of metaphors and heuristic frameworks across multiple disciplines and domains. In this regard, space has served as a long-standing metaphor and model in philosophy for conceiving and expressing abstract relations among concepts. For instance, the philosopher Wilfrid Sellars developed an original critique of certain forms of empiricism by organizing his thought broadly in terms of a distinction between two conceptual "spaces": the *natural space of causes* and the *logical space of reasons* [15]. In using this spatial metaphor to mark a conceptual difference, Sellars stands in a long line of philosophical articulation of conceptual and logical relations in terms of spatiality, whether geometrical (distance, curvature, congruence) or topological (continuity, covering, inclusion), that goes back at least to Aristotle and the Stoics with their use of the spatial concept of "topos" as a way to express common conceptual and rhetorical "places". In recent decades, this age-old metaphorization of relations among concepts as spatial relations has taken a more rigorous mathematical turn with the category-theoretic notion of "topoi" and, more generally, the concentrated investigation of structures—mutually illuminative—common to logic and topology [1, 10, 17].

In this chapter we will show how some of the elementary mathematical tools of category theory provide a useful framework for formalizing the intuition of a conceptual space within which the "movements" or "gestures" of abductive inference may be localized and tracked. Category theory itself is a highly general area of Mathematics that originally emerged within algebraic topology (and thus is naturally adept for abstract representations of general "spaces") and yet also lends itself readily to furnishing formal models for epistemic situations and abductive inferences, among many other applications [5, 16]. Here, we present a general model of the concepts and dynamics of *communities* of agents and practical *obstructions* varying over logical "spaces" of pragmatic *commitments*. The basic tool from category theory that we use to integrate these three main components of the model is called a *presheaf*, which is essentially a mathematical coordination of abstract relational structures (categories) with concrete realizations of those structures in systems of sets and functions. Within this framework of presheaves, the epistemic and conjectural situations that characterize abductive inference appear as subclasses of systematic triadic coordinations of commitments, communities and obstructions. The advantage of using category theory to model these systematic relations is that it abstracts away from unnecessary concrete details of the objects under investigation in whatever context in order to examine only the system of relations (which may often be conceived of naturally as structure-preserving mappings or transformations) structuring the objects themselves *as* representatives of some particular type. In the case of modeling abductive inference in terms of varying communities and obstructions relative to a space of commitments, the use of category theory promises to be helpful in understanding abductive reasoning by representing and analyzing the relevant dynamics at an appropriately abstract level of relationality.

The core idea in the present approach involves considering abductive inferences as movements within spaces constituted by communities of agents defined by their epistemic and conjectural commitments. The contents of such commitments are themselves understood as determined by the possible events or states that might obstruct them. The framework outlined below provides a way to characterize the anomalies or "surprising events" that give rise initially to an abductive inference and also to tease apart three parameters or dimensions of the abductive "movement" itself through conceptual space.

The categorical approach outlined below depends upon the combination or, better, coordination of two guiding ideas, one primarily mathematical and the other primarily philosophical:

1. The guiding mathematical idea is to use categories of *Heyting algebras* as the formal theoretical framework for modeling and investigating the systematic logical relationships among commitments. Heyting algebras are useful for examining systematic part/whole relationships in a variety of contexts, in particular in ways that bring to light deep connections between logic and topology [17]. Categories of Heyting algebras will serve as the abstract "spaces of commitments" through which abduction "moves".
2. The guiding philosophical idea involves tracking processes of reasoning in terms of pragmatic relations among communities of agents and among practical obstructions that structurally mirror those of systems of commitments. There is a natural dialectic whereby communities relate to one another "in accord with" logical entailments of commitments (covariance) and collections of obstructions are gathered "oppositely" (contravariance). On the one hand, we thus consider communities in a given population as organized pragmatically via the commitments that they do or do not share. On the other hand, we conceive of processes of reasoning as organized through a semantic coordination of such pragmatic commitments and objective obstructions that represent material incompatibilities with respect to those commitments. This general pragmatic-semantic approach follows the lead of [3].

We show in what follows how category theory, and in particular presheaves, can integrate these two basic ideas in a straightforward way and how abductive reasoning may be modeled in that formal context.

2 Affectivity and Abduction

It is important, first, to emphasize some of the difficulties that stand in the way of any project committed to formalizing abduction. Unlike the cases of deductive and inductive inference, when aiming for a formal model of abductive inference it may not be entirely clear what the intuitive concept that one is trying to formalize even *is*. Working notions of abduction range from very general conceptions that assimilate even ordinary perception to the schema of abductive inference to very specific conceptions of abduction as inference to the best explanation (IBE) within some definite theoretical context. At the theoretical level, no single framework encompasses all these approaches, although some proposals have gained traction in local contexts (for example [2, 11] in the context of belief revision expressed in formal languages). The study of abduction remains, to speak in Kuhnian terms, pre-paradigmatic. We do not presume to offer a framework that would serve as such a paradigm, but by taking a new tack on this old problem we hope to contribute to the ongoing conversation and the healthy proliferation of points of view out of which a paradigm might eventually emerge.

Discussions of abduction typically trace both the word itself and the tradition that stems from it to Peirce [7, 13]. We too take our cue from Peirce's overall approach to the problem and begin from the phenomenological experience of abductive inference, rather than presuming a formal context at the outset. In this regard, Peirce uses distinctive and idiosyncratic language to formulate his concept of abductive inference. We present one of his canonical formulations here, separating and enumerating the argument in three parts and highlighting in italics his more colorful turns of phrase ([13] 5.180–212 in [14] vol. 2, p. 231, numerals and emphases added):

1. The *surprising* fact, C, is observed;
2. But if A were true, C would be a matter of course.
3. Hence, there is reason *to suspect* that A is true.

Notice that the phrases highlighted in (1) and (3) employ directly *affective* language. Surprise and suspicion are affective states. To be sure, they may and usually do possess a cognitive component as well, but such a component is not essential to them. At least in the case of surprise, one may very well experience the subjective feeling of surprise before one has conceptually grasped any determinate object to be surprised *at*. What makes surprise and suspicion what they are is a more or less definite kind of feeling, a sort of existential status or subjective mood, that is, a determinate *affect*. If abduction can only be adequately characterized in terms of such affective states, then attempts to formalize abduction in purely cognitive or propositional terms are bound, at least in part, to miss the mark.

The affects of surprise and conjectural suspicion depend upon enduring subjective and pragmatic dispositions that are themselves expressed in complex webs of expectation and attention. One is surprised, for instance, when one's ordinary expectations are overturned. Often such expectations remain more or less unconscious, as when one indifferently "expects" the status quo. Conjectural suspicion tends to be more conscious and focused, but it too can only be understood in terms of dispositions and expectations that both shift and endure over time. If abduction crucially depends upon such affective states and pragmatic dispositions, then formal models of abduction must start from conceptual frameworks distinct from those appropriate to static propositions or beliefs. The language and conceptual architecture of "commitments" appears promising in this regard. Commitments, like affects, are complex. They can involve both conscious and unconscious elements as well as both subjective and objective determinations. Commitments themselves are not necessarily affective in a direct manner (it is not clear, for instance, how the commitment to following traffic laws *feels*), but due to the typical coordination of subjective and objective poles of commitment, there is a natural way in which commitments, to the extent that they can be blocked, assisted, shared, and so forth, become linked in regular and tractable ways to certain affective states of the agents who hold those commitments.

Such complexity cannot be adequately captured by the coarse-grained structure of atomic propositions and classical Boolean values, but perhaps it *can* be represented to some reasonable degree of faithfulness by the more flexible determinations of continuities and overlaps of open sets in an appropriate topological space. In line with this consideration, we present a general pragmatic framework based on the notion of an abstract *space of commitments* in which the shades and nuances of the affective dynamics of abductive inference may be characterized in terms of the expression of commitments across various *communities* on the one hand and the potential *obstructions* that might arise to those commitments on the other.

A helpful intuition to guide understanding of the formal construction below may be to think of the space of commitments as organized somewhat like a movable archery target superposed over some given territory. To possess some degree of commitment to the realization of X relative to various similar and different possibilities is, metaphorically speaking, to target a region X with some degree of precision with respect to its neighboring regions. Commitments come in degrees, like concentric circles on the target, although some commitments are incommensurable, as when the target is moved to a different (non-overlapping) area. Relaxing a commitment corresponds to enlarging the target zone X. The more relaxed the commitment, the easier it then is to hit the target (because the target region is larger). Similarly, sharpened or more specialized commitments correspond to restricting the size of the target region. Hitting a bullseye is hard because its area of success is so small relative to the surrounding area of failure. The space of commitments is in this way like a territory transcribed with a variety of target circles of various sizes, with some separate, some overlapping, some wholly included in others. A movement through the space of commitments corresponds to a change from one selected target circle to another, that is, the relaxation, strengthening or redirecting of the region that counts as a successful hit. The context-bound significance of the various targeted regions will be determined by how their successful targeting might be blocked or obstructed by various factors. And collectivities or communities will be determined on the basis of shared acts of projective targeting.

3 The Formal Framework

The overarching intuition for the following framework is that of a logical space of commitments, with possible obstructions and actual communities varying over that space in a reasonable and controlled way. In other words, it may be helpful—in accordance with the "target map" metaphor outlined above—to think of any particular target region on the map (that is, some particular commitment of whatever degree of generality or specificity) as defined on the one hand by a collection of possible obstacles to attaining it that thereby determine its "meaning" and manifested on the other hand by a collection of agents who aim at it and thereby concretely realize the targeting commitment in whatever context. As the target region varies through the available space (expanding, contracting, shifting position, etc.)—which corresponds to controlled changes in features of the specified commitment—the coordinated collections of obstructions and communities of agents linked to that commitment will, in the proposed model, vary accordingly.

In the following sections, we will first explain how the space of commitments itself may be formalized within category theory as a category in which commitments are the objects and entailment relations between commitments are the arrows or morphisms. When such a category meets certain specifications (as outlined below), it will constitute a *Heyting algebra* and may therefore be understood as the abstract representation of a topological space. We then organize a kind of "semantics" for these categories of commitments by associating each commitment in the category with a set of obstructions to that commitment. These are required to respect the entailment relations in an appropriate way. Finally, we characterize the realizations of categories of commitments by populations that are parceled out into various communities according to the commitments they share. As with obstructions, the formation of such communities is required to appropriately respect the entailment relations between the commitments in the underlying space.

3.1 Commitments

At the basis of the formal model is a space of commitments which arises from a class of commitments taken together with certain prescribed relations between pairs of those commitments. We conceive of these relations as arrows going from one commitment to another. In this way, the logical space of commitments appears as a directed graph. Given two commitments, say c_1 and c_2, we have the relation $c_1 \longrightarrow c_2$ if and only if commitment c_1 entails commitment c_2. A typical class of such implications will be generalizations: for instance, commitment to reading *War and Peace* entails commitment to reading something written by Tolstoy, which entails commitment to reading something by either Tolstoy or Dostoevsky. In accordance with the categorical approach, we refrain from specifying this entailment relation in itself (considerably simplifying our task!) and attend only to the structural behavior of this relation insofar as it forms a unified system via its specifiable meta-relations. This will constitute the system of the logical space of commitments itself as a category and, indeed, a Heyting algebra.

We require the following characteristics to ensure the structural integrity of the space of commitments. Each requirement is subject to a natural and plausible interpretation in terms of our ordinary intuitions of how commitments are related to one another[1]:

1. Each commitment entails itself.
2. Commitment entailment is transitive.
3. Given two commitments, c_1 and c_2, we have the conjunctive commitment $c_1 \wedge c_2$ such that any commitment that entails both c_1 and c_2 also entails $c_1 \wedge c_2$.
4. Given two commitments, c_1 and c_2, we have the disjunctive commitment $c_1 \vee c_2$ such that any commitment that is entailed by both c_1 and c_2 is also entailed by $c_1 \vee c_2$.
5. Given two commitments, c_1 and c_2, we have the implicative commitment $c_1 \Rightarrow c_2$ such that for all commitments x where $c_1 \wedge x$ entails c_2, x entails $c_1 \Rightarrow c_2$.
6. There is a "self-contradictory" or "false" commitment $\perp$ such that $\perp$ entails every other commitment.
7. There is a "maximally general" or "always true" commitment $\top$ such that any commitment entails $\top$.

For simplicity's sake, we also assume that two commitments that mutually imply one another "collapse" to form a single commitment.[2]

A note of clarification is needed for requirement 5: the commitment $c_1 \Rightarrow c_2$ can be thought of as the element that represents the "most general implication" of its type in the following sense. Consider the collection of all the conjunctive commitments "generated" by a commitment p (that is, of the form $p \wedge a$ for some a), restrict this collection to those that entail q and call this restricted collection $X(p, q)$. Then $p \Rightarrow q$ is the least general commitment overall such that it is entailed by every element of $X(p, q)$. Thus, $p \Rightarrow q$ is the *strongest* commitment (the least general) with the property that it is entailed by all of $X(p, q)$ (and is therefore more general than each of them).

Taken together, these structural requirements on a system of commitments organized by a relation of entailment induce the formal structure of a category and, indeed, a Heyting algebra. Of course, this is no accident. We have simply expressed (one formulation of) the familiar Heyting algebra axioms in terms of commitments and their entailment-relations.[3] This way of proceeding thus uses an abstract mathemat-

---

[1]We omit the corresponding formal characterizations for simplicity of presentation.

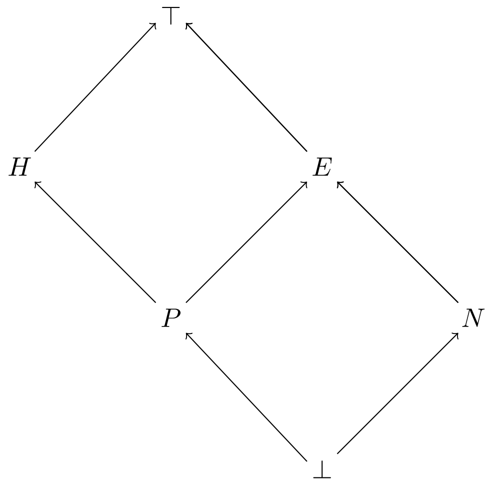
[2]From a logical point of view, this assumption amounts to conceiving of the commitment-space as a Lindenbaum algebra.

[3]Formally, a Heyting algebra is defined as follows [6]: First, we define a *bounded lattice* as a partially ordered set $\mathcal{A}$ in which for every pair $a, b \in A = \mathrm{Ob}(\mathcal{A})$ there exists a *supremum* or *least upper bound* $a \vee b$ and an *infimum* or *greatest lower bound* $a \wedge b$ and there exists a *top* and a *bottom*, 1 and 0, respectively, such that for every $a \in A$, $a \preceq 1$ and $0 \preceq a$. In categorical terms, the supremum of a and b is their coproduct $a + b$, while the infimum is $a \times b$, and the bottom is an *initial object*

ical characterization of space (at once logical and topological) to make a precise "spatial metaphor" for systems of subjective commitment. Such an approach will be justified to the extent that it will allow for the further construction on its basis of a more complex theoretical framework that may itself be useful in understanding abduction.

The diagram below represents a concrete example. It will serve as something of a "toy model" to exhibit the dynamics of abductive inference in this framework. For simplicity's sake, in our example we will consider only *theoretical* commitments. This has the advantage of fitting into the familiar conception of abductive inference quite readily. It has the disadvantage, however, of working within a rather narrow class of commitments. It should be borne in mind throughout the exposition below that the dynamics that are characterized in this particular theoretical context are special cases of a more general class of dynamics within commitments of whatever type.

From a strictly formal point of view, the diagram shows a lattice of nodes bound by an maximal node $\top$ and a minimal node $\bot$. Between these upper and lower bounds are four nodes labeled H, P, E and N. P is below H and E, and N is below E (when the arrows are thought of as an order-relation). We omit the identity arrows that belong to each object (each node) and it should be remembered that every path of arrows also defines an arrow because of transitivity.



such that for every object a, there exists a unique morphism $0 \to a$. Analogously, the top is a *terminal object* such that there exists a unique $a \to 1$ for every object a. A *Heyting algebra* is then a bounded lattice $\mathcal{A}$ with two functors $\wedge : A \times A \to A$ and $\Rightarrow: A \times A \to A$, left and right adjoints, respectively. The interpretation is that given $c, a, b \in A$, $x \wedge a \preceq b$ is equivalent to $x \preceq a \Rightarrow b$. In logical terms, $x \wedge a$ is a *conjunction* between x and a, while $a \Rightarrow b$ is an *implication* with antecedent a and consequent b. In categorical terms $a \Rightarrow b$ is the exponential b^a.

This toy example may be taken to picture Rutherford's experimentally-based critique of the "plum-pudding" model of the atom and the scientific community's shift of commitment from that model to a model of atoms with a nucleus [12]. We consider the space of available theoretical commitments as composed of four regions, each of which corresponds to a general hypothesis: the homogeneity of matter (H), the existence of electrons (E), the atomic plum-pudding model (P), and the nucleated atomic model (N). We presume that commitment to P entails commitment to H and E and that commitment to N entails commitment to E.[4] It should be emphasized that this is no more than a toy model used for expository purposes and is by no means meant to capture the real complexity of the epistemic issues at stake in this historical example.

3.2 Obstructions

The Heyting algebra category organizes the entailment relations among the commitments at stake, which correspond to inclusions of open sets in an abstract topological space. This system of relations is somewhat like the skeleton of a formal syntax. It tells us only how the commitments in the system are linked to one another via entailment relations, but it tells us nothing of how they should be understood in themselves. There is no "semantics" (in a rough and intuitive sense) for this abstract "syntactical" space. We will provide such a semantics by coordinating sets of obstructions with each commitment in the space.

Truth is the adequation of intellect and thing. Philosophers have tied themselves into knots trying to clarify just what this relation of "adequation" is and how to test or verify it. The project of empiricism presumes a sphere of subjective experiences characterized by their various qualities and an objective world of things with various properties. Truth occurs when the qualities match up appropriately with the properties. But this "matching up" which seems at first glance so straightforward turns out to raise a whole host of problems that have occupied philosophy at least since Hume.

One proposal for getting around these problems amounts to shifting the terrain from subjective representations to subjective commitments. The basis of the link between mind and world thus shifts from "matching up" to "being compatible with". Truth is then lifted to a higher degree of abstraction. It no longer consists of a relation between a thought and a thing, but is instead conceived as a relation between a system of relations between thoughts and a system of possibilities among things. This program was proposed and developed by Wilfrid Sellars in the 1960s and 1970s

[4] It should be noted that this particular interpretation serves a merely heuristic purpose. The structures and dynamics discussed in the following sections do not depend upon any particular interpretation of the commitments in the Heyting algebra presented, only upon the structure of the algebra (category) itself.

[15] and carried forward more recently by Robert Brandom. It is a research program characterized, as the title of [4] has it, by the shift *from empiricism to expressivism*. It may equally be described as a shift from representationalism to pragmatism. Part of its power comes from specifying a general *practical* or *pragmatic* context of correlation between subjective commitments and objective conditions and then positioning traditional philosophical problems surrounding truth and knowledge within this more general context in terms of specifically *epistemic* or *alethic* commitments. In our study of abduction we will be concerned primarily with communities, not individuals, but for purposes of exposition it will be simpler here to consider just individual commitments. In the following section we will broaden this notion to include communities.

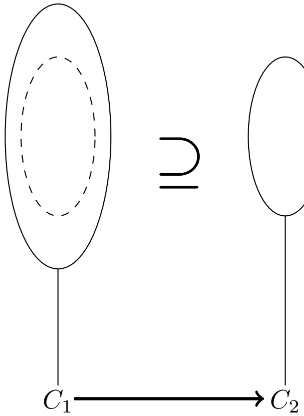
The core of this program involves the following two-level insight: subjective commitments are correlated with objective conditions of material compatibility and incompatibility (level one), and inferences from certain commitments to certain other commitments correspond (or *should* correspond) to relations between their correlated conditions of material compatibility and incompatibility (level two). At level one, for instance, my commitment to buying an ice cream cone is correlated on the one hand with a variety of possible flavor preferences, different weather conditions, various locations, various social settings, etc. (material compatibilities) and on the other hand with conditions that would block or invalidate such a commitment: lacking sufficient funds, no ice cream in stock, and so forth (material incompatibilities).

The role of materially *incompatible* conditions is especially important here. It is in general impossible to specify all the conditions that are compatible with a given commitment. The world is, after all, a complicated business, and indeterminately specified global diversity is its default state. But it is quite often possible to specify if not *all* the conditions that are incompatible with a given commitment, at least a robust and relatively complete set of local conditions any one of which would be sufficient to obstruct or compromise that commitment. For example, incompatible (more or less completely) with the commitment to buying ice cream is (a) not wanting ice cream, (b) not having enough money to buy ice cream, and (c) not having ice cream available for sale.

Following roughly in the path of the Spinozist insight *omnis determinatio est negatio* ("all determination is negation"), correlating subjective practical commitments with sets of incompatible objective conditions effectively *defines* those commitments in terms of what would exclude them or make them unrealizable.

This lays the ground for the second level of coordinating commitments and conditions. It is clear on the one hand that certain commitments entail certain other commitments. For instance, the commitment to watch a film in a movie theater entails (under normal convivial conditions) the commitment to avoid conversing loudly with one's neighbor. It is clear on the other hand that if such entailment relations (and the inferences they underwrite) are to be non-arbitrary, they must be grounded in the objective conditions correlated with the commitments they link together. In fact, there is quite natural way to coordinate the entailment relations among commitments on the one hand and the relations among sets of objective conditions on the other. When a commitment C_1 properly entails some other commitment C_2, it is to be expected that the set of objective conditions incompatible with C_2 should be a subset of the conditions incompatible with C_1. In other words, if C_2 follows from C_1, then everything incompatible with C_2 should be incompatible with C_1.

We thus have a general situation that may be pictured in the following way:



Let us unpack what the diagram expresses. There are essentially two layers: a base layer and an upper layer. The base or lower layer consists of two nodes or "objects" representing distinct commitments and an arrow linking them. This arrow is meant to represent a relation of entailment that holds between the two commitments: commitment to C_1 is understood to entail commitment to C_2. Each of the objects in the lower layer is linked to a oval "fiber" in the upper layer. Each of these fibers is intended to represent the set of objective conditions that are incompatible with the commitment linked to it below. The key insight is that if indeed C_1 entails C_2 then any element of the set of objective conditions incompatible with C_2 (an element of the fiber over C_2) should also be incompatible with C_1 (and thus an element of the fiber over C_1). This is because anything that obstructs C_2 or makes it impossible will also by that very fact (given the entailment relation) make C_1 impossible. If commitment C_1 can be maintained under conditions that make C_2 impossible, then clearly C_1 does not necessarily entail C_2. The reasoning here is, in effect, a pragmatic instantiation of the logical schema of *modus tollens*.

Only two terms appear in each of the layers in the given diagram. But it should be noted that the diagram could be extended under the same interpretation and it would preserve the indicated structure transitively in both the base layer and the fiber layer. In other words, if the base layer were extended by one node so as to express $C_1 \longrightarrow C_2 \longrightarrow C_3$, this would imply $C_1 \longrightarrow C_3$ by transitivity of implication. The new base node C_3 would be equipped with its own fiber of incompatibilities. And if we designate the fibers over C_1,C_2 and C_3 as $F(C_1)$,$F(C_2)$ and $F(C_3)$ respectively, then in the upper layer we would (by the same reasoning given above) have the transitive sequence of inclusion relations $F(C_1) \supseteq F(C_2) \supseteq F(C_3)$.

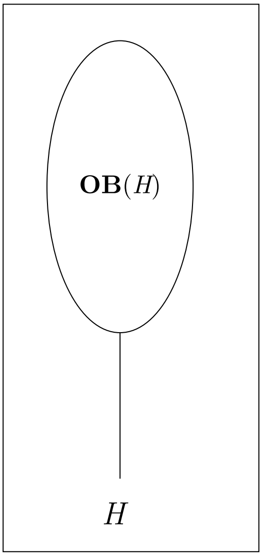
So a relation of commitment-entailment in the base layer moving from left to right (in the diagram) is coordinated with a relation of incompatibility-inclusion in the upper layer moving from right to left. The important point to note is not the particular orientation in the given diagram (left to right in the base; right to left in the fibers) but rather the intrinsic opposition in the two directions relative to one another, the fact that the relations go in opposite directions. This coordinated opposition of the direction of one system of relations with respect to some other system of relations is called, in a formal register, *contravariance*. In the present context, this systematic "oppositional parallelism" expresses in a highly general way just how the obstructive semantics is coordinated with the spatial (topological) syntax. It is the special character of category theory to exhibit such general mappings of systems of relations in an immediate and easily formulated manner.

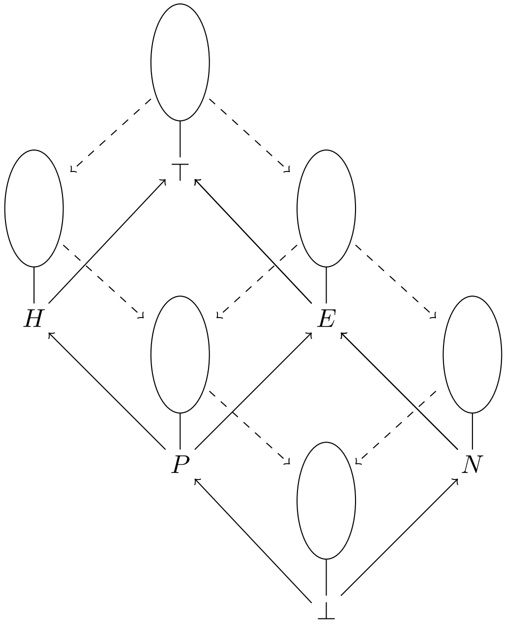
The point is not at all that pragmatic commitments and sets of incompatible conditions can be perfectly specified and exhaustively determined in any given case. They almost certainly cannot, save for the most trivial cases. What matters is the form of the relation between the implicative relations in the lower or base layer and the inclusion relations in the upper or fiber layer. Category theory is especially useful in modeling such relations because it abstracts away from isolated determinations of individuals and attends only to relational and systematic determinations.

We are now prepared to combine (a) the idea that there is a natural contravariance between subjective commitments and objective incompatibilities and (b) the mathematical structure of Heyting algebras as a way to organize relations among communities of commitment together in a single integrated framework. The basic construction we will use is a fundamental tool within category theory called a *presheaf* [8, 9].

We take the example of Rutherford provided above and consider a contravariant functor from that category into the category $\mathbf{Sets}_{\subseteq}$.[5] We call this functor **OB**. To each commitment (object) c in the space of commitments (category) $\mathcal{C}$, we assign the set $\mathbf{OB}(c)$ of possible obstructions to c, represented by an oval on top of c. The dashed arrows represent inclusion functions between sets (note the contravariance).

[5] $\mathbf{Sets}_{\subseteq}$ notates the category whose objects are sets and whose arrows or morphisms are all inclusion functions between sets. This category is a faithful subcategory of the category **Set** of sets and functions.





It is unnecessary to fill in all the details of this particular mapping, since from a categorical perspective what is relevant is only that some such mapping has been specified. We attend not to the details of what is going on "under the hood" but instead only to their effects as manifest in the external relations in the specified categories. Somewhat as a computer programmer typically works in a relatively high-level programming language and does not need in general to attend to the underlying machine code, a theorist of abductive inference may (once the underlying structural interaction of the commitment space and the obstructions presheaf over it is understood) effectively take this structure for granted as a well-behaved domain for conceiving, expressing and testing various abductive dynamics. In order to motivate the exhibition of abductive inferential dynamics below in terms of our fixed toy model, we will simply note here that among the obstructions to commitment to the hypothesis of the "plum-pudding" atomic model (node P in the base category) we may posit experimental results sufficiently similar to those found by Rutherford and his team. If we label such an experimental obstruction η, then we have the situation that $\eta \in \mathbf{OB}(P)$ and $\{\eta\} \subseteq \mathbf{OB}(P)$.

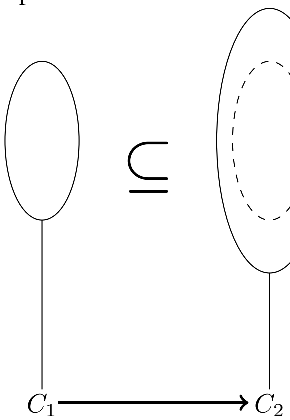
In accordance with the insight noted above, we want to specify that arrows between the sets of obstructions represent inclusions of sets. This will ensure the plausibility of our semantics. But this specification comes for free by choosing the category $\mathbf{Sets}_{\subseteq}$ as our target category. The mathematical infrastructure of category theory ensures the very coordination of commitment-space "syntax" and "semantics" that is our aim.

3.3 Communities

The coordination of the syntactical space of commitments with the semantic determinations of sets of obstructions lays out a system of potential meaningful commitments in some given context. But *who or what* exactly would hold and thus manifest these commitments? To realize this system of possible commitments as a collection of actual determinations in a concrete world requires specifying individuals or groups who would in fact hold these commitments, thus realizing them concretely *as* actual subjective commitments. We add this layer of determinations in a way that is parallel to the procedure followed in determining the semantics of obstructions, but which is organized in the systematic "direction" opposite to it.

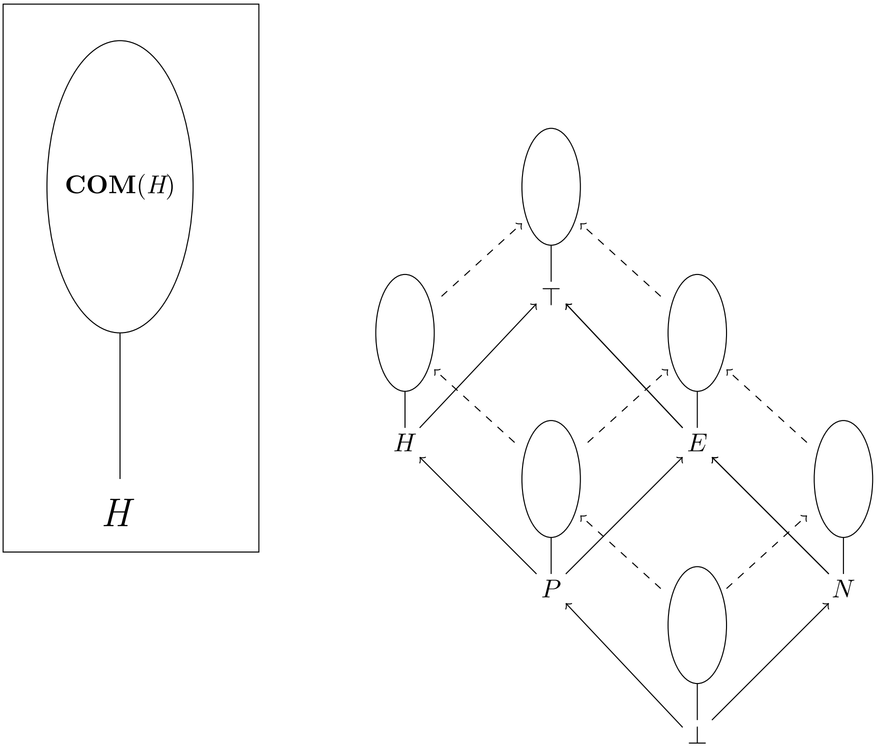
Given a space of commitments, we may consider some population of agents distributed across that space. To each commitment in the space, we then associate the collection of agents in the population who actually hold that commitment. In order to respect the implication-relations among the commitments in the space, it seems natural to require that any agent who is committed to some commitment c_1 must also be committed to c_2 if $c_1 \longrightarrow c_2$. Of course in real-world contexts, agents are—alas—not always logically consistent and may be committed, for instance, to fostering high-quality public education without at the same time being committed to funding such education. The purpose of specifying the framework in this way is not to exclude such possibilities, but only to prescribe an initial "ideal case" in such a way that different (and perhaps more realistic) scenarios may be specified relative to this ideal model.

We thus have a situation parallel but inverse to that of obstructions:



To each commitment in a commitment space, we associate a community composed of agents who share that commitment. If a commitment implies another commitment, then the community associated with the former is necessarily committed to the latter. Thus we have a system of inclusion relations among sets that runs parallel to the implication arrows in any given commitment space. This situation is that of a covariant functor from the commitment space (understood as a category) into the category of $\mathbf{Sets}_{\subseteq}$ of sets and inclusions.

We continue with the example of Rutherford provided above and consider a covariant functor from that category of theoretical commitments into the category $\mathbf{Sets}_{\subseteq}$, having first fixed some set P to be our target population set of relevant scientists. We call this functor **COM**. To each commitment (object) c in the space of commitments (category) $\mathcal{C}$, we assign the set $\mathbf{COM}(c)$ of individuals in P committed to c. The dashed arrows represent inclusion functions between sets (note the covariance).



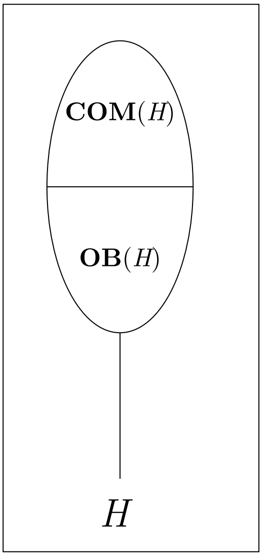
Just as with obstructions to commitments, we do not need to specify exactly what the constituent elements of these sets and the details of these mappings are. For present purposes it is enough to specify that $\mathbf{COM}(P)$ and $\mathbf{COM}(N)$ are non-empty. By the very fact that the meet $P \wedge N$ is $\bot$ we infer that (in this context) simultaneous commitment to both these hypotheses is impossible and they thus represent competing research programs. Thus, $\mathbf{COM}(\bot)$ is the empty set.

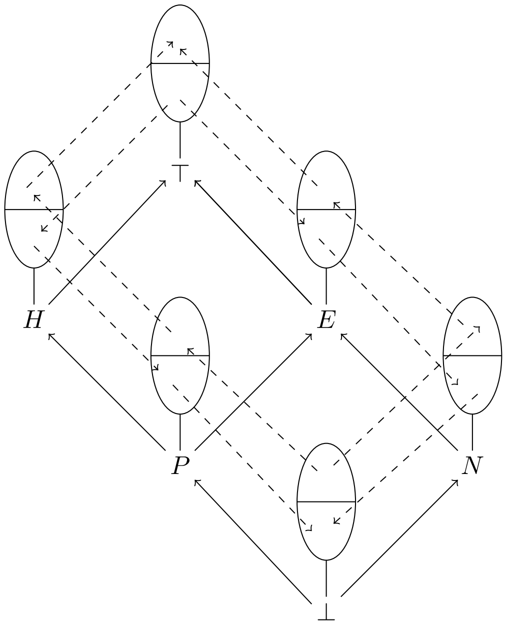
The formal definition of a conceptual "space" coordinating commitments, obstructions and communities that will serve as a context for abduction may now be given. A *context for abduction* is a Heyting algebra category $\mathcal{C}$ equipped with a pair of functors, namely, a contravariant functor $\mathbf{OB} : \mathcal{C}^{\mathbf{op}} \longrightarrow \mathbf{Set}_{\subseteq}$ and a covariant functor $\mathbf{COM} : \mathcal{C} \longrightarrow \mathbf{Set}_{\subseteq}$ such that for all c objects of $\mathcal{C}$ $\mathbf{COM}(c) \cap \mathbf{COM}(\neg c) = \varnothing$.[6] This relatively compact definition captures a surprisingly robust collection of fea-

[6]Here, as is standard, $\neg c$ is shorthand for $c \Rightarrow \bot$.

tures that would be expected of populations of reasonable agents sharing various commitments that are determined by their entailment-relations on the one hand and their possible material obstructions on the other. Category theory allows for such straightforward modeling of systems of structural relations, and here provides a mathematically rigorous (if highly idealized) topological setting—a precise spatial metaphor, so to speak—for framing problems and solutions to the theorization of abductive inference.

Combining the diagrams that are intended to express the two functors **OB** and **COM** into a single integrated diagram, we may picture the sets **OB**(c) and **COM**(c) associated with some commitment c as the lower and upper layers, respectively, of a single two-layer fiber. Functions between the sets in the lower layers are contravariant, and functions between sets in the upper layers are covariant. The contravariant functor **OB** and the covariant functor **COM** are thus "stacked" one on top of the other. This does not entail any additional mathematical structure or formalism; it is merely a convenient visual representation of the two functors considered simultaneously. Continuing with the case developed above, we thus have the following informal but suggestive picture:





This diagram pictures the structural coordination of (semantic) contravariant obstructions and (actualized) covariant communities with respect to the space of (potential) commitments defining a particular theoretical context. The different ways this diagram could be “filled in” by definite functors **OB** and **COM** correspond to the various possible concrete contexts of commitment that might define a given timeslice.[7]

4 Abductive Inferences Within Spaces of Commitments

At last, within the categorical framework we have taken pains to develop it becomes possible to characterize the structural dynamics of abductive inference. We proceed in two main steps which correspond to the two subjective affects noted above in Peirce’s description of abduction. First, we characterize the “surprising event” that occasions an abductive inference in terms of one or more of the obstructions in $\mathbf{OB}(c)$ for some given commitment c. Then, we characterize the reasonable suspicion of an abductive hypothesis in terms of three distinct parameters: the redistribution of committed communities, the reconfiguration of obstructions coordinated with particular commitments, and the refinement of the commitment space itself. The dynamics of an abductive inference may involve modifying any combination of one or more of these parameters. In all such cases, the categorical framework preserves overall coherence in the space as a whole as these changes are made.

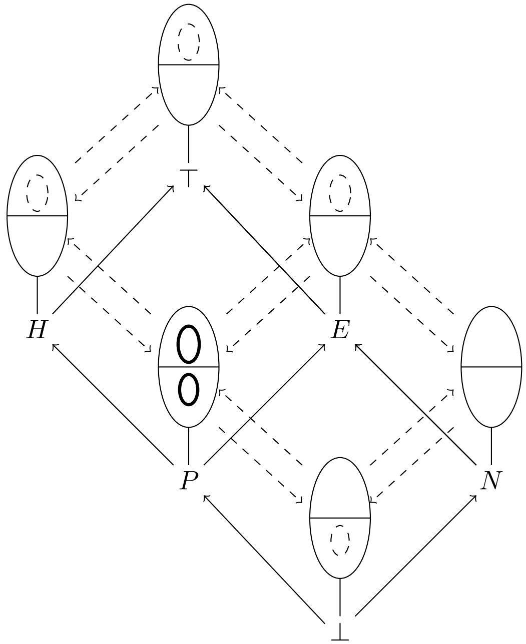
4.1 The Occasion for Abduction: Surprise

First of all, we may characterize the occasion or instigation for abduction in a straightforward manner. Some part of a committed community encounters one or more obstacles to that commitment. Peirce’s “surprising event” in the present framework will involve two terms, a set of agents and a set of obstructions. A surprising event, from this point of view, is just a subset of the community committed to some commitment who encounter a subset of the obstacles that obstruct that commitment.

Formally, given a context for abduction, an obstructed commitment or “surprise” is a triple $< c, OB_c, COM_c >$ consisting of some commitment $c \in \mathcal{C}$ and sets $OB_c \subseteq \mathbf{OB}(c)$ and $COM_c \subseteq \mathbf{COM}(c)$.

---

[7]Of significant interest, but beyond the reach of the present exposition, is the system of controlled variation among such concrete timeslices. This system may be formalized in terms of the category of *natural transformations* among the relevant pairs of **OB** and **COM** functors.



In this diagram, the lower of the two thick ovals in the fiber over P represents the obstruction η to the plum-pudding model posed by Rutherford's experimental data. The upper oval represents a collection of scientists committed to that model who encounter Rutherford's data. In the formal framework, these are represented by a subset of $\mathbf{OB}(P)$ and a subset of $\mathbf{COM}(P)$ respectively. Let us designate these as $\{\eta\}$ (as in Sect. 3.2) and $\{\sigma_1, \sigma_2, \sigma_3\}$ (assuming three "surprised" plum-pudding scientists: σ_1, σ_2, and σ_3). Thus, we have $\{\eta\} \subseteq \mathbf{OB}(P)$ and $\{\sigma_1, \sigma_2, \sigma_3\} \subseteq \mathbf{COM}(P)$. Intuitively, the encounter of these three scientists with the data obstructing their commitment to the plum-pudding model are, first of all, surprised to find their theoretical commitment challenged in this way and, secondly, impelled to revise their theoretical commitment in light of this obstruction. We will track this latter process of theoretical revision in the subsequent section.

The dashed ovals in the diagram represent the subsets $\{\sigma_1, \sigma_2, \sigma_3\} \subseteq \mathbf{COM}(H)$, $\{\sigma_1, \sigma_2, \sigma_3\} \subseteq \mathbf{COM}(E)$, and $\{\sigma_1, \sigma_2, \sigma_3\} \subseteq \mathbf{COM}(\top)$ on the one hand and the subset $\{\eta\} \subseteq \mathbf{OB}(\bot)$ on the other that are mandated by the structures of the respective presheaves. It is important to remember that the coherence of the model as a whole depends upon these covariant and contravariant subset inclusions mirroring the structure of the underlying category. Local modifications in the fibers (for instance over the node P) will for this reason have ramifications elsewhere in the model. Thus, local changes will typically induce global updates in the presheaf functors.

4.2 *The Dynamics of Abduction: Reason to Suspect*

The core of abductive inference is a certain kind of subjective response made in the wake of a surprising event. In particular, an abductive inference requires the additional ingredient of some *motivation* to understand or explain the source of the surprise. This motivational element and its relation to the explanatory power of the abduced conjecture mark an especially tricky aspect of the overall problem of modeling and formalizing abduction. We will return to this problem below. For now, we only suggest that the motivation to explain—which is not reducible to the affect of surprise but must be adjoined to it—is, at least in principle, describable in terms of additional modes, or loci, of commitment.

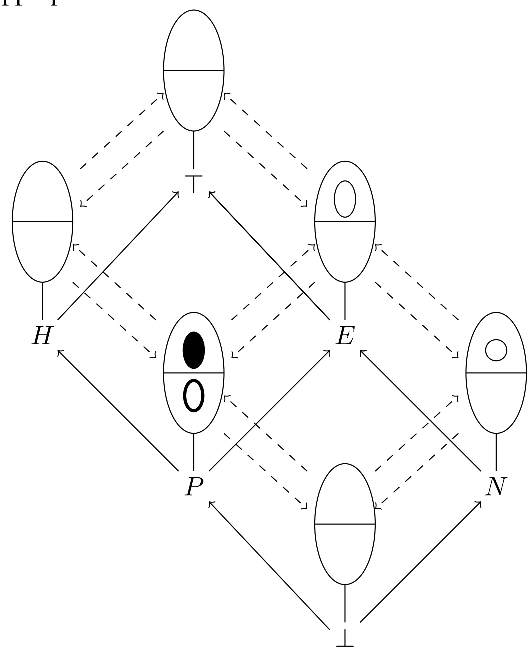
The "spatial" model developed above has three basic aspects: commitments, obstructions and communities. Each of these aspects is structured by a system of relations, and the three systems of relations are coordinated with one another by way of the categorical machinery of presheaves. Accordingly, when conceiving of abductive inference within this model, three types of movement or change through its "space" suggest themselves as potential components of abductive revision.[8] The communities may be redistributed; the obstructions may be reconfigured; the commitment space itself may be refined. The structural dynamics of these three processes are sketched below. We propose that a robust theory of abduction should take into account the interactions and mutual constraints among all three of these coordinated aspects.

4.2.1 Redistribution of Communities with Respect to Commitments

In the current framework, we conceive of an abductive inference as first and foremost an updating of what agents in some global context hold which commitments. To make an abduction is, typically, to give up certain old commitments and to take on certain new ones. In a community-based context, this means splitting up or emptying certain communities of commitment and forming new communities or forging alliances with those already there. From the current formal point of view, this amounts to redistributing the communities throughout the space of commitments, that is, updating the functor **COM**. Thus, the first and most important parameter for abductive inference is the redistribution of communities of commitment in the face of some given surprise.

[8]The ancient Greek concept of $\kappa\iota\nu\eta\sigma\iota\sigma$ would hold resources here for understanding and expressing the direct connection between movement and change, even to the point of identifying them.

The subset of agents in a community who encounter an obstruction to their common commitment are led to revise that commitment in light of the obstruction. The typical shape of such revision is that of an initial relaxation of their commitment followed by a subsequent resharpening or refocusing of commitment onto an alternate goal or project. To return to the "target" metaphor used above, a reasonable response to some commitment obstruction consists of increasing the relevant target zone (so that a near miss may now count as a hit) and then tightening the target circle to exclude the obstruction (which now falls outside the target zone). In the more abstract categorical Heyting representation, this movement appears as an "ascent" from the associated object in the category in the direction of $\top$, that is, "following the arrows" of the category until the relevant obstruction is no longer present in the associated fiber, followed by a "descent" that goes "against the arrows" to some more specialized commitment that resolves or ameliorates the obstruction in whatever way is contextually appropriate.

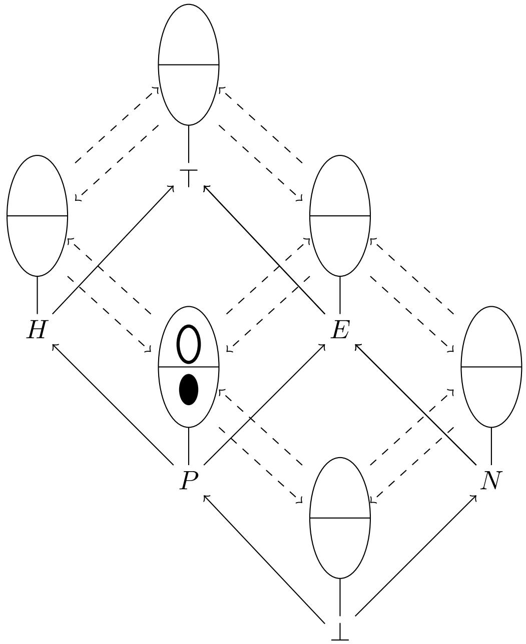


The full technical details of the formalization of this "movement" go beyond the scope of this presentation, but a simple example should illustrate the main idea. In the given case of competing atomic hypotheses, the subset of the community committed to the plum-pudding atomic model who have encountered the obstructive data in Rutherford's work cannot remain committed to that hypothesis and must "move" elsewhere (that is, change their commitments). The most reasonable approach is to move "up" the lattice until the obstructive data is, precisely, no longer obstructive. In this case, the data also obstructs the more general hypothesis that matter is homogeneous ($\{\eta\} \subseteq \mathbf{OB}(H)$), so the path to the left from P would have to go all the way up to $\top$, a disappointing result. The path to the right, however, leads to commitment E, the hypothesis that electrons exist and $\{\eta\} \nsubseteq \mathbf{OB}(E)$. Let us assume that σ_1 and σ_2 do no more than this. Thus, they merely relax their commitment to the plum-pudding model (P) so that they are now committed to no more than the hypothesis E (that electrons exist). Scientist σ_3, however, performs the more interesting further "descent" from E to N. Commitment to the hypothesis N not only escapes the obstruction η; it also, if it is true, *would explain* η. Thus, σ_3 appears to have performed an abduction in a significantly stronger sense than the mere "commitment revision" of σ_1 and σ_2.

The movements of σ_1, σ_2 and σ_3 through the space of commitments are represented by updating the functor **COM** to a new functor **COM**$'$. In the updated functor **COM**$'$ that "replaces" **COM** in light of the surprising encounter with obstruction η, $\{\sigma_1, \sigma_2\} \subseteq \mathbf{COM}'(E)$ and $\{\sigma_3\} \subseteq \mathbf{COM}'(N)$. All other fibers are, to the greatest extent possible, preserved.

4.2.2 Reconfiguration of Obstructions with Respect to Commitments

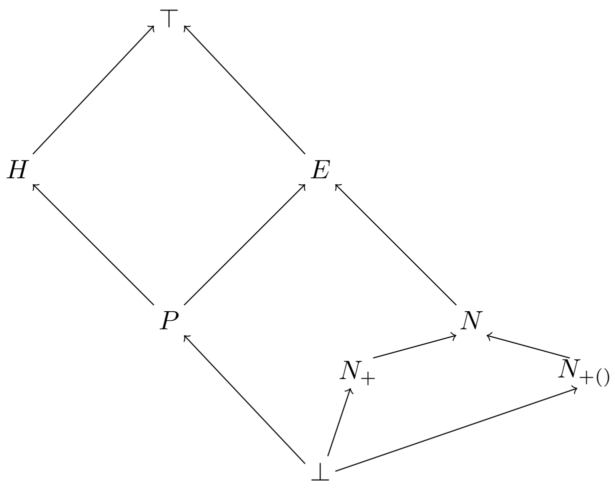
In the proposed categorical framework, sets of obstructions provide a pragmatic semantics for the commitments in the base category by varying over them in a controlled way in the fibers of the functor **OB**. In response to an encounter with one or more of these obstructions, agents may revise their commitments as outlined above in a way that tracks the global semantics in a reasonable manner. But it is also possible to hold the commitments and associated communities fixed and to revise the semantics instead. Trivially, it is possible simply to "delete" the relevant obstruction(s) from the fiber of the given commitment. Scientists committed to the plum-pudding model may decide by fiat, for example, that Rutherford's experimental data does not represent an obstruction to their commitment. In deciding this, they are in effect redescribing the content of their commitment. Note that such a modification of the semantics of the commitment space (the fibers in the **OB** presheaf) would seem to recommend revisions or refinements of the commitment space itself. This suggests that a more thorough model would need to elaborate and formalize the complex interplay between the formal entailment structure in the base category and the semantic content codified by the sets of obstructions and their inclusion relations.



In the diagram above, the obstruction η has been deleted from the set $\mathbf{OB}(P)$. Thus, a new functor $\mathbf{OB}'$ has been induced such that $\mathbf{OB}'(P) = \mathbf{OB}(P) - \{\eta\}$. Generally, such a deletion would entail corresponding changes to fibers upstream if they too contain the obstruction η. In order to maintain the integrity of the presheaf, local semantic changes involve more general regional effects. In the given example, we may assume that Rutherford's experimental results also obstruct commitment to the hypothesis H that matter is homogeneous ($\eta \in \mathbf{OB}(H)$). Consequently, the update $\mathbf{OB}'(H) = \mathbf{OB}(H) - \{\eta\}$ is "forced" by the deletion of η from $\mathbf{OB}(P)$. Once again, the particularly abductive character of the change remains largely implicit in this simple example. In order for the abductive inference not only to *accommodate* but (potentially) to *explain* the surprising event, it would seem that the reason for denying or deleting the relevant obstruction must in some way be incorporated into the new commitment. This key aspect remains to be brought explicitly into the dynamics of the model.

4.2.3 Refinement of the Commitment Space

Finally, abductive inference may involve modifying the Heyting algebra category $\mathcal{C}$ that represents the underlying abstract space of commitments. In particular, the conceptual distinctions marked out by the objects and relations constituting the space of commitments may be *refined* so that more nuanced adjustments can be registered. Formally, this means passing from the category $\mathcal{C}$ to a new category $\mathcal{C}'$ such that there is a surjective Heyting algebra homomorphism (a functor) $h : \mathcal{C}' \longrightarrow \mathcal{C}$. Since the **OB** and **COM** functors are defined with respect to $\mathcal{C}$, any change in $\mathcal{C}$ forces corresponding updates in **OB** and **COM**, even if the functors are intended to be preserved as much as possible. Roughly speaking, the sets in the fibers over the newly introduced objects of the category must distribute the sets in the fibers from the original functor **OB** (**COM**) immediately above (below) them.[9]



Here, the movement of abductive inference is understood not as a passage from one pre-given commitment-state to another, but as the enrichment or differentiation of the current epistemic commitment-state. In this case, Rutherford's nuclear hypothesis represented by the node N in the Heyting category is differentiated into two sub-hypotheses N_+ (the atomic nucleus consists only of protons) and $N_{+()}$ (the atomic nucleus consists of both protons and neutrons). Clearly, commitment to either N_+ or $N_{+()}$ entails commitment to N, and neither N_+ or $N_{+()}$ entails the other. Note that the set of objects $\{\bot, N_+, N_{+()}, N\}$ constitutes a sub-Heyting algebra with the arrows inherited from the larger category in which they are embedded. In general, the refinement of a single commitment c in a space of commitments $\mathcal{C}$ will consist of a expansion of the object c into a sub-Heyting algebra in which c is the local $\top$ or $\bot$ (together with all the concomitant relations necessitated by the categorical structure itself: transitivity as composition, etc.). This formal refinement at the categorical level nicely captures the intuitive idea of elaborating the "inner logic" (the space of internal distinctions and entailments) of a given commitment.

[9] We omit the technical details for reasons of space.

5 Conclusion

It should be clear that the dynamics outlined above do not provide anything like a complete formal theory of abductive inference. In particular, the core notion of conjectured *explanation* remains insufficiently articulated. What has been laid out does offer, however, a collection of general "shapes" within the structural dynamics that may be thought of metaphorically as characteristic gestures expressing different aspects of abduction. But the delicate fine-tuning necessary to chart out all the possible nuances and interconnections of these gestures awaits further research. Nevertheless, within this still relatively coarse framework, those aspects of abductive inference that have shown themselves to be most resistant to formalization may at the very least be characterized—precisely *as problems*—in new and potentially productive ways. For instance, the explanatory aspect of a conjectured hypothesis may be conceived (roughly) as its incorporation of a commitment (properly qualified as conjectural) to some claim h such that if h is true then some relevant aspect of the instigating obstruction would be understood as "a matter of course". In this way, it seems that the specific explanatory character of abductive inference will consist in drawing appropriate conditional antecedents for the relevant subset of obstructive instigation into the space of commitments itself. In combination with the other parameters of adjustment available within this framework, it seems that various new ways to posit general and specific types of abduction as well as to generate diverse classes of concrete models of such types become possible for future research programs.

Abductive inference is a highly complex natural process, a reasoning activity of which only sophisticated cognitive agents are capable. Because of the essential integration of abductive inference within the practically-oriented and affectively-tinged experiences of subjective agents in dynamic material contexts, formal theories of abduction need to be responsive to multiple dimensions of subjectivity, objectivity and controlled variability. The framework outlined here coordinates three such dimensions: logical spaces of subjective commitments, potential objective obstructions to those commitments, and communities realized and organized with respect to their common commitments. It should be clear that the actual process of abductive inference cannot be reduced to any one of these dimensions. A robust theory of abduction would need to synthesize the three modes of redistribution, reconfiguration and refinement sketched above into a unified account that would at once preserve their relative independence and also guarantee their coherent composition. Such a thorough account cannot be provided here. But it may be hoped that the utility of category theory and the particular tool of presheaf functors over "spaces" (Heyting categories) for constructing theoretical frameworks that track both the independence and the interdependence of distinct dimensions of a problem in a controlled manner has been sufficiently demonstrated.

References

1. Bell, J.L.: Toposes and Local Set Theories: An Introduction. Clarendon/Oxford University Press, Oxford (1988)
2. Boutilier, C., Becher, V.: Abduction as Belief Revision. Artif. Intell. **77**, 43–94 (1995)
3. Brandom, R.: Articulating Reasons: An Introduction to Inferentialism. Harvard University Press, Cambridge, MA (2000)
4. Brandom, R.: From Empiricism to Expressivism: Brandom Reads Sellars. Harvard University Press, Cambridge, MA (2015)
5. Caterina, G., Gangle, R.: Iconicity and Abduction. Springer, Berlin (2016)
6. Esakia, L.: Heyting Algebra: Duality Theory. Springer, Basel (2019)
7. Fann, K.T.: Peirce's Theory of Abduction. Martinus Nijhoff, The Hague (1970)
8. Fong, B., Spivak, D.I.: An Invitation to Applied Category Theory: Seven Sketches in Compositionality. Cambridge University Press, Cambridge, UK (2019)
9. Gangle, R.: Diagrammatic Immanence. Edinburgh University Press, Edinburgh (2015)
10. Goldblatt, R.: Topoi: The Categorical Analysis of Logic. Elsevier, Amsterdam (1984)
11. Harman, G.: The Inference to the Best Explanation. Philos. Rev. **74**, 88–95 (1965)
12. Lakhtakia, A.: Models and Modelers of Hydrogen. World Scientific, Singapore (1996)
13. Peirce, C.S.: Hartshorne, C., Weiss, P., Burks, A. (eds.) Collected Papers, vols. 1–8. The Belknap Press of Harvard University Press, Cambridge (1931–1958)
14. Peirce, C.S.: Peirce Edition Project. The Essential Peirce, vols. 1–2. Indiana University Press, Bloomington (1998)
15. Sellars, W.: Philosophy and the scientific image of man. In: Science, Perception and Reality. Ridgeview Publishing Co., Atascadero, CA (1962)
16. Spivak, D.I.: Category Theory for the Sciences. MIT Press, Cambridge, MA (2014)
17. Vickers, S.: Topology via Logic. Cambridge University Press, Cambridge, UK (1989)

Developing and Applying Peirce's Abduction

Abduction as “Leading Away”

Aristotle, Peirce, and the Importance of Eco-Cognitive Openness and Situatedness

Lorenzo Magnani

Abstract In this article I will take advantage of the logical and cognitive studies I have illustrated in my recent book *The Abductive Structure of Scientific Creativity. An Essay on the Ecology of Cognition* (2017), in which the process of building new hypotheses is clarified thanks to my *eco-cognitive model* (EC-Model) of abduction. Also resorting to a new interpretation of Aristotle’s seminal work on abduction, I will emphasize the crucial role played in abductive cognition by the so-called “optimization of eco-cognitive openness and situatedness”. Indeed, in the chapter B25 of *Prior Analytics* concerning ἀπαγωγή (“leading away”), we can see some of the current well-known distinctive characters of abductive cognition already expressed, which are in tune with the EC-Model: Aristotle is still pointing to the fundamental inferential role in reasoning of those externalities that substantiate the process of “leading away” (ἀπαγωγή). Hence, we can gain a new positive perspective about the “constitutive” eco-cognitive character of abduction, just thanks to Aristotle himself. Situatedness is related to eco-cognitive aspects: to favor the solution of the abductive problem input and output of the formula

$$\Lambda_1, \ldots, \Lambda_i, ?_I \Vdash^X_L \Upsilon_1, \ldots, .\Upsilon_j$$

have to be thought as optimally positioned: indeed I also contend than a disregarded issue concerning abduction is related to the current lack of knowledge about what I call “discoverability” and “diagnosticability”. In the formula above $\Vdash^X_L$ indicates that inputs and outputs do not stand each other in an expected relation and that the modification of the inputs $?_I$ can provide the *abductive solution*. In general, in this characterization the direction is not from evidence/premises to abductive result but the forward fashion is adopted, where the inferential parameter $\Vdash$ sets some appropriate logical relationship between an input which consists in both the abductive guess to be found and a background theory (or just some premisses), and an output—for example an evidence, a novel phenomenon to be abductively “explained” through facts, rules,

L. Magnani (✉)
Department of Humanities, Philosophy Section and Computational Philosophy Laboratory, University of Pavia, Pavia, Italy
e-mail: lmagnani@unipv.it

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_4

or even new theories. Further, in the case of scientific settings, this optimality is made possible by a maximization of changeability of both input and output: not only inputs have to be enriched with the possible solution but, to do that, other inputs have usually to be changed and/or modified. This changeability first of all refers to a wide epistemological openness.

1 Eco-Cognitive Model of Abduction (EC-Model), Aristotle's Ἀπαγωγή as "Leading Away", and Its Eco-Cognitive Openness

At the center of my perspective on cognition is the emphasis on the "practical agent", of the individual agent operating "on the ground", that is, 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, artifacts, 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, that is in reasoning to hypotheses.

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" ([39], 5.283). It is clear that this semiotic view is considerably compatible with my perspective on cognitive systems as embodied and distributed systems. It is in this perspective that we can fully appreciate the role of abductive cognition, which not only refers to propositional aspects but it is also performed in a framework of distributed cognition, in which also models, artifacts, internal and external representations, manipulations play an important role.

The backbone of this approach can be found in the manifesto of my eco-cognitive model (EC-model) of abduction in [24].[1] 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 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

[1]Further details concerning the EC-model of abduction can be found in [27, 28].

esthetic abductions (referring to creativity in art, literature, music, games, 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[2] in order to become the "best" choice: 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 judgements: 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 "can" become prescriptive and "true," informing the agent's behavior 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 selective or creative abduction[3] with their related cognitive strategies 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!

Aristotle presents a seminal perspective on abduction, which is in tune with my EC-Model: indeed Aristotle's abduction (ἀπαγωγή)[4] exhibits a clear eco-cognitive openness. In the chapter B25 of *Prior Analytics* concerning ἀπαγωγή ("leading away"), a chapter also studied by Peirce, we can see some of the current well-known distinctive characters of abductive cognition already expressed, which are in tune with the EC-Model. By providing an illustration of the role of the method of analysis and of the middle terms in Plato's dialectic argumentation, considered as related to the diorismic/poristic process in ancient geometry—also, later on, emphasized by Proclus—I maintain that it is just this intellectual heritage which informs Aristotle' chapter B25 on ἀπαγωγή. Even if, in general, Aristotle seems to sterilize, thanks to the invention of syllogistic theory, every "dialectic" background of reasoning, nevertheless in chapter B25 of *Prior Analytics* he is still pointing to the fundamental inferential role in reasoning of those externalities that substantiate the process of

---

[2]These kinds of abduction can be called "knowledge enhancing": Peirce implicity provides various justifications of the knowledge enhancing role of abduction, that is when abduction is not considered an inference to the best explanation in the classical sense of the expression, that is an inference necessarily characterized by an empirical evaluation phase, or inductive phase. In chapter three of [29] I have provided the example of conventions: abducing *conventions* favors 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]For example, selective abduction is active in diagnostic reasoning, where it is merely seen as an activity of "selecting" from an encyclopedia of pre-stored hypotheses; creative abduction instead refers to the building of new hypotheses. I have proposed the dichotomic distinction between selective and creative abduction in [22]. A recent and clear analysis of this dichotomy and of other classifications emphasizing different aspects of abduction I have described is given in [37].

[4]In this article for every English word or expression that refers to Plato and Aristotle's texts from *Meno*, *Phaedo*, and *Prior Analytics* I have indicated the original Greek in parentheses. Since the translations are not uniformly satisfactory, it is best to include the original Greek.

"leading away" (ἀπαγωγή). Hence, we can gain a new positive perspective about the "constitutive" eco-cognitive character of abduction, just thanks to Aristotle himself.

First of all, we have to take note that it seems Peirce was not satisfied with the possible Apellicon's correction of Aristotle's text about abduction: "Indeed, I suppose that the three [abduction, induction, deduction] were given by Aristotle in the *Prior Analytics*, although the unfortunate illegibility of a single word in his MS, and its replacement by a wrong word by his first editor, the 'stupid' [Apellicon],[5] has completely altered the sense of the chapter on Abduction. At any rate, even if my conjecture is wrong, and the text must stand as it is, still Aristotle, in that chapter on Abduction, was even in that case evidently groping for that mode of inference which I call by the otherwise quite useless name of Abduction—a word which is only employed in logic to translate the ἀπαγωγή of that chapter" ([39], 5.144–145, *Harvard Lectures on Pragmatism*, 1903).

At this point I invite the reader to carefully follow Aristotle's chapter from the *Prior Analytics*, studied and cited by Peirce, I will soon quote and analyze in detail in the following Sect. 1.1. Aristotle's discussion turns arguments that transmit the uncertainty of the minor premiss to the conclusion, rather than the certainty of the major premiss. If we regard uncertainty as an epistemic property, then it is reasonably sound also to say that this transmission can be effected by truth-preserving arguments: by the way, it has to be said that this is not at all shared by the overall Peirce's view on abduction, which is not considered as truth preserving.

I want first of all to alert the reader that in the case of the Aristotelian chapter, abduction does not have to be discussed keeping in mind the schema of the fallacy of affirming the consequent. What is at stake is abduction considered either (1) the classification of a certain "unclear" dynamic argument in a *context-free* sequence of three propositions; or (2) the introduction in a similar "unclear" dynamic three-propositions argument (in this case no longer *context-free*) of few new middle terms. Hence, ἀπαγωγή— that the translator of the *Prior Analytics* I am adopting usefully renders with "leading away" (abduction)—is, exactly (in the Aristotelian words we will soon entirely report below)

1. the feature of an argument in which "it is clear (δῆλον) that the first term belongs to the middle and unclear (ἄδηλον) that the middle belongs to the third, though nevertheless equally convincing (πιστόν) as the conclusion, or more so" ([1], B25, 69a, 20–22, p. 100);
2. the introduction of suitable middle terms able to make the argument capable of guiding reasoning to substantiate an already available conclusion in a more plausible way: Aristotle says in this way we "are closer to scientific understanding": "if the middles between the last term and the middle are few (ὀλίγα) (for in all these ways it happens that we are closer to scientific understanding (πάντως γάρ ἐγγύτερον εἶναι συμβαίνει τῆς ἐπιστήμης)]" ([1], B25, 69a, 22–24, p. 100).

[5] Apellicon was the ancient editor of Aristotle's works. Amazingly, Peirce considers him, in other passages from his writings, "stupid" but also "blundering" and "scamp" ([20], p. 248).

It is clear that the first aspect indicates a certain status of the uncertainty of the minor premiss and of the conclusion and of the related argument (in abductive terms, we can say that the conclusion represents a very provisional hypothesis that has to be rendered plausible); the second aspect, from the perspective of the eco-cognitive model of abduction, directly refers to the need, so to speak, of "additional/external" interventions in reasoning, exactly to a get a better and more acceptable conclusion/hypothesis. It has to be said that Aristotle does not consider the case of the creative reaching of a *new* conclusion/hypothesis (that is of a creative abductive reasoning): however, I have illustrated in [27] that this last case appears evident if we consider the method of analysis in ancient geometry, as a mathematical argument which mirrors the propositional argument given by Aristotle, provided we conceive it in the following way: *we do not know the conclusion/hypothesis, but we aim at finding one thanks to the tentative introduction of further "few" suitable middle terms*.

I cannot emphasize enough, as I have anticipated above, that in this Aristotelian chapter B25—I will soon report—we face the influence of the method of analysis and of the middle terms in Plato's dialectic argumentation, and the fundamental role of diorismic/poristic processes in ancient geometry (see below in this article), which constitute the main aspects of the use of *hypothetical argument* in ancient philosophy [21]. I think that disregarding this heritage renders difficult a full appreciation of Aristotle's chapter B25. Indeed underestimating these geometric and dialectic roots, and considering Peircean conception of abduction as only referred to the fallacy of affirming the consequent, in a kind of logicocentric reduction of Peircean richer analysis of abduction, also characterized by philosophical, psychological, cognitive, and semiotic considerations, is extremely reductive. This reduction leads recent examinations of Aristotle's passage to divergent—and at the same time traditional [20]—conclusions [5, 12]. These skeptical conclusions contend that in chapter B25 we are only and basically dealing with Aristotelian dialectic, not with abduction, and Peirce made a mistake (he himself shows some hesitation regarding his reading of Aristotle) forcing the interpretation of chapter B25 by considering it at the roots of the modern concept of abduction. I think these conclusions misrepresent what Aristotle is doing in this chapter of *Prior Analytics*. On the contrary I think that it is exactly the exploitation of the dialectic background that makes available the new perspective on abduction as a kind, Woods says, of "third-way reasoning": in the following paragraphs and in Sect. 1.4, in which I will further explain the relationships between dialectical and syllogistic logic, I will motivate my point of view.

Indeed, the current revival of studies on abduction can further benefit from recent explicit Woods' claim about the need of a *naturalized* logic of the so-called *third-way reasoning*, I have just quoted above,[6] which refers to those kinds of inference that owe their "rightness to the meeting of standards other than deductive validity and inductive strength", certainly overworked and overvalued by logicians, and that instead show the capacity to reach good and reliable results. Third-way reasoning refers to those cases of human reasoning such as abduction, but also to many other kinds of the so-called fallacies, disregarded or simply considered irremediable errors

[6] On this issue cf. also the recent [26, 31].

by logic since its ancient and modern beginnings. The study of third-way reasoning aims at going beyond the logical obsession for *consequence-having*, opening to the analysis of the structure of the so-called *consequence-drawing* (eventually truth-preserving or truth-generating), typical of various actual human performances ([51], pp. 3, 24, 293, and 518). Woods' recovering of the cognitive and inferential positive value of fallacies basically concerns: *ad hominem*, *ad populum*, *ad verecundiam*, *ad ignorantiam*, affirming the consequent (abduction), denying the antecedent, begging the question (and circularity), many questions, hasty generalization, equivocation, gambler's fallacy, base rate, and *post hoc, ergo propter hoc* (including false cause).

We know that dialectic is a process for discovering the truth about some subject thanks to questions conceived to elicit what answerers already know about it implicitly or by exposing contradictions and confusions in the answerer's position: what is important to note here is that, as Woods contends: "Whatever else it is, a dialectical logic is a logic of consequence-drawing", that is not a logic of "consequence-having" ([49], p. 31). Woods further emphasizes that "The logic of syllogisms, he [Aristotle] insisted, was a logic of consequence-having. Dialectic, he said, was (among other things) a theory of consequence-drawing.[7] But the target properties of syllogistic logic would be wholly definable without reference to properties needed for the description of attack-and-defend encounters between human individuals" ([51], p. 31).[8] In sum, no surprise to see Aristotle to illustrate the first uncertain features of the logic of abduction, contrasted with the main character of valid syllogisms,[9] instead immunized with respect to external interferences, thanks to a reference to Plato and geometry.

Aristotle himself expressly contends that the *necessity* of valid syllogism is related to the circumstance that "no further term from outside (ἔξωϑεν) is needed", in sum not only Aristotle aims at banishing redundant premisses from syllogisms, but consequently indicates that syllogism is the fruit of a kind of eco-cognitive *immunization* [valid syllogism occurs in the emptiness, so to speak, of a kind of "logical space", Woods says ([51], p. 24)]. Woods contends that an Aristotle's great metalogical achievement lies in the proof (given in *Posterior Analytics*) that all the truths of a mature deductive science lie in the demonstrative closure of its first principles, and

---

[7]Recently Woods has changed a bit his position: there are also the consequence-spotting arguments, for example the refutation-arguments of *Sophistical Refutations*, the refuted party cannot add to his beliefs the conclusion of the refutation since it contradicts the thesis he's been defending. Consequence-having occurs in logical space, consequence-spotting occurs in psychological space, and consequence-drawing occurs in the inferential subspace of psychological space ([54], p. 14). Woods also contends that the syllogistic is the logic of consequence-having only for syllogisms-as-such; for syllogisms-in-use, consequence-having remains a background condition, but the specific targets of the arguments are not themselves satisfaction of the consequence-having relation.

[8]It is interesting to address the reader to a recent research in the field of logic, which approaches abduction without disregarding its pragmatic/dialectical dimension, thanks to the adoption of a dialogical logic. This current logical illustration of the dialectic involved in abduction is able to model argumentative interactions leading to conjectures. In this case the authors can conclude: "Thus, the consequence-having is dialectified and the having-drawing distinction is somewhat broken" [3].

[9]Aristotle insists that all syllogisms are valid (by definition) [52], there is no such thing as an invalid syllogism. We know the syllogistic tradition began to relax this requirement quite early on.

do so in a way that produces the knowledge of their truth. Here, too, there is no hint of dialectical taint. Aristotle says "A deduction (συλλογισμὸς) is a discourse (λόγος) in which, certain things having been supposed, something different from the things supposed results of necessity because these things are so. By 'because these things are so', I mean 'resulting through them,' and by 'resulting through them' I mean 'needing no further term from outside (ἔξωθεν) in order for the necessity to come about'" ([1] A1 24, 20–25, p. 2).

1.1 Aristotle Chapter B25 of the **Prior Analytics**

The following is the celebrated chapter B25 of the *Prior Analytics* concerning abduction. The translator usefully avoids the use of the common English word *reduction* (for ἀπαγωγή): some confusion in the literature, also remarked by Otte ([36], p. 131), derives from the fact reduction is often rigidly referred to the hypothetical deductive reasoning called *reductio ad absurdum*, unrelated to abduction, at least if intended in Peircean sense. Indeed, the translator chooses, as I have anticipated, the bewitching expression "leading away".

> XXV. It is leading away (ἀπαγωγή) when it is clear (δῆλον) that the first term belongs to the middle and unclear (ἄδηλον) that the middle belongs to the third, though nevertheless equally convincing (πιστόν) as the conclusion, or more so; or, next, if the middles between the last term and the middle are few (ὀλίγα) [for in all these ways it happens that we are closer to scientific understanding (πάντως γάρ ἐγγύτερον εῖναι συμβαίνει τῆς ἐπιστήμης)]. For example, let A be teachable, B stand for science [otherwise translated as "knowledge"], and C justice [otherwise translated as "virtue"]. That science is teachable, then, is obvious, but it is unclear whether virtue is a science.[10] If, therefore, BC is equally convincing (πιστόν) as AC, or more so, it is a leading away (ἀπαγωγή) (for it is closer to scientific understanding (ἐγγύτερον γάρ τοῦ ἐπίστασθαι) because of taking something in addition (τὸ προσειληφέναι), as we previously did not have scientific understanding (ἐπιστήμη) of AC). Or next, it is leading away (ἀπαγωγή) if the middle terms between B and C are few (ὀλίγα) [for in this way also it is closer to scientific understanding (εἰδέναι)]. For instance, if D should be "to be squared," E stands for rectilinear figure, F stands for circle. If there should only be one middle term of E and F, to wit, for a rectilinear figure together with lunes to become equal to a circle, then it would be close to knowing (ἐγγύς άν εἴη τοῦ εἰδέναι). But when BC is not more convincing (πιστότερον) than AC and the middles are not few (ὀλίγα) either, then I do not call it leading away (ἀπαγωγή). And neither when BC is unmiddled: for this sort of case is scientific understanding (ἐπιστήμη) ([1] B25, 69a, 20–36, pp. 100–101).

This passage is very complicated and difficult, also because, as ([19], p. 25) says "Aristotle transforms (as usually) the argument in a syllogistic form". I have indicated words and expressions in ancient Greek because they stress, better than in English, some of the received distinctive characters of abductive cognition:

---

[10]We have to anticipate (see below Sect. 1.3.2) that, having given the geometrical description of the argument "from hypothesis" Plato/Socrates, in the *Meno*, had already clearly considered the statement "Is virtue teachable?" a genuine hypothesis: "In the same way with regard to our question about virtue, since we do not know either what it is or what kind of thing it may be, we had best make use of a hypothesis in considering whether it can be taught or not" ([44], 87b).

1. ἄδηλον [unclear] refers to the lack of clarity we are dealing with in this kind of reasoning; furthermore, it is manifest that we face a situation of ignorance—something is not known—to be solved. This type of reasoning

   > […] might be said, semi-demonstrative, semi-dialectical, since it has a minor premiss which is known, and a minor premiss which is uncertain. The unknown character of the minor premiss is the essential feature of this kind of syllogism, according to Aristotle (see *An. pr.*, 69a35–36) and this uncertainty is transferred to the conclusion. Now, the minor premiss is getting "more known" (69a28) and the conclusion more probable, if we introduce a middle term which links the subject "virtue" to the predicate "knowledge" ([19], pp. 25–26);

2. πιστόν [convincing, credible] indicates that degrees of uncertainty pervade a great part of the argumentation;
3. the expression "then it would be close to knowing (ἐγγὺς ἀν εἴη τοῦ εἰδέναι)", which indicates the end of the conclusion of the syllogism,[11] clearly relates to the fact we can only reach credible/plausible results and not ἐπιστήμη itself; Peirce will say, similarly, that abduction reaches plausible results and/or that is "akin to the truth"[12];
4. the adjective ὀλίγα [few] dominates the passage: for example, Aristotle says, by referring to the hypotheses/terms that have to be added—thanks to the process of leading away—to the syllogism: "Or next, it is leading away (ἀπαγωγή) if the middle terms between B and C are few (ὀλίγα) (for in this way also it is closer to scientific understanding (εἰδέναι))". The term ὀλίγα certainly resonates with the insistence on "minimality" that dominates the first received models of abduction of the last decades of XX century.

I favor the following interpretation ([41], p. 173): abduction denotes "the method of argument whereby in order to explain an obscure or ungrounded proposition one can lead the argument away from the subject to one more readily acceptable".

In the passage above Aristotle gives the example of the three terms "science" [knowledge], "is teachable", and "justice" [virtue], to exhibit that justice [virtue] is

[11] Let me reiterate that Aristotle insists that all syllogisms are valid; there is no such thing as an invalid syllogism. The syllogistic tradition began to relax this requirement: here I will use the term syllogism in this modern not strictly Aristotelian sense.

[12] I agree with Pietarinen recent limpid analysis that attributes to Peirce the seminal refinement of the ancient concept of abduction in terms of the precise moods and modalities involved in conjecture-making [42]: "The mood of abductive conclusions is not only interrogative or imperative but rather a complex mixture of them and closer to what linguistics call co-hortative or jussive moods; those that capture both the important idea of 'pursuit-worthiness' of abduced conclusions as well as the 'rational hopes' of our guesses to turn out in the way our minds or machines predict. Investigands are invitations to proceed investigating conjectures further. But those investigations have to start off at the level of pre-beliefs. It is in the nature of the logic of abduction that some reasons are found why its conclusions are worthy of further investment. Abductive conclusions cannot be neutral indicative or epistemic statements; they carry normative and pragmatic force. […] we see Peirce taking abductive conclusions drawn in an interrogative mood, considered as a formalization of the Socratic questioning method". I think these considerations and this conclusion about Socratic method further support my referral to the important role of Plato's dialectics in the background of Aristotle's considerations about abduction as "leading away", I am presenting in this article.

teachable: Aristotle is able to conclude that justice [virtue], is teachable, on the basis of an abductive reasoning, that is ἀπαγωγή, endowed with a certain plausibility. We have to reiterate that, having given the geometrical description of the argument “from hypothesis” Plato/Socrates clearly considers the statement “Is virtue teachable?” a genuine hypothesis: “In the same way with regard to our question about virtue, since we do not know either what it is or what kind of thing it may be, we had best make use of a hypothesis in considering whether it can be taught or not” ([44], 87b) (see below Sect. 1.3.2). A second example of *leading away* is also presented, which illustrates that in order to make a rectilinear figure equal to a circle only one additional middle term is required.

I do not think appropriate to consider, following Kraus ([20], p. 247), the adumbrated syllogism (first Aristotelian example in the passage above)

AB Whatever is knowledge, can be taught
BC Virtue (e.g., justice) is knowledge
AC Therefore virtue can be taught.

just an example of a valid deduction, so insinuating Peirce’s interpretation failure. I think Peirce is crystal clear about the issue, presenting this case—even if with doubts and remarks—as a typical way of considering abduction, as I will better explain below in Sect. 1.3.1.

Indeed, it seems vacuous to elaborate on the syntactic structure of the involved syllogism, as Kraus does: the problem of abduction in chapter B25 is embedded in the activity of the inferential mechanism of “leading away” performed thanks to the introduction of new terms, as I explained above. He also says that the second Aristotelian example

Whatever is rectilinear, can be squared
A circle can be transformed into a rectilinear figure by the intermediate of lunes
Therefore, a circle can be squared

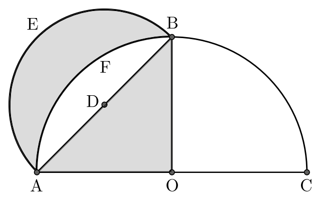
still a simple deduction, was questionably supposed by Peirce to be fruit of the correction of Aristotle’s original text due to the “stupid” Apellicon, considered responsible of blurring Aristotle’s reference to abduction. Indeed, Kraus suggests that, following Peirce, the original text would have to be the following:

Whatever is equal to a constructible rectilinear figure, is equal to a sum of lunes
The circle is equal to a sum of lunes
Therefore, the circle is equal to a constructible rectilinear figure.

which indeed fits the Peircean abductive schema.

It is true that Peirce, following what I called his “sentential” perspective on abduction ([24], chapter two) is certainly inclined, in his considerations of these two examples, to think that some Apellicon’s bad corrections of the Aristotle’s text would have concealed the form—instead correct—indicated by Kraus and reported above. This figure of the syllogism, which notoriously illustrates the fallacy of affirming the consequent, is certainly privileged by Peirce in his sentential view of abduction, but we have to stress that he also adds some comments which testify his doubts: “If

Fig. 1 Hippocrates' result: if the external arc of the lune is a semicircle, and the internal one is a quarter of circle, then the two grey areas are equal and the lune is squarable



we do not suppose this corruption to have taken place, we are reduced to accepting the text as it stands; and if we accept the text as it stands we must accept the usual interpretation of it. This interpretation is that abduction is nothing but an ordinary syllogism of the first figure, when we are not sure of the minor premiss, but still are more inclined to admit it than we should be to admit the conclusion if the latter were not a necessary consequence of the former" ([39], 7.251). Moreover, as already reported in Sect. 1, Peirce is also sure that Aristotle is "evidently groping for that mode of inference which I call [...] Abduction": "At any rate, even if my conjecture is wrong, and the text must stand as it is, still Aristotle, in that chapter on Abduction, was even in that case evidently groping for that mode of inference which I call by the otherwise quite useless name of Abduction—a word which is only employed in logic to translate the ἀπαγωγή of that chapter" ([39], 5.144–145, *Harvard Lectures on Pragmatism*, 1903). In sum, Aristotle is actually establishing the first building blocks of the concept of abduction, by referring to the use of hypothetical argument in Plato and pre-Euclidean geometry.

Let us come back to the second—geometrical—example, which opened a wide-ranging and controversial debate in the literature on Peirce's abduction. Actually Hippocrates of Chios proved that the lune bounded by the arcs labeled E and F in Fig. 1. has the same area as triangle ABO. This afforded some hope of solving the circle-squaring problem, since the lune is bounded only by arcs of circles. As Odifreddi explains ([35], p. 55), to prove that a circle is squared it would be sufficient to prove that the lune whose external arc is a semicircle and the internal arc is one sixth of a circle is squared. Unfortunately, as proved by Nikolai Chebutarjov in 1934 and Arkadiy Dorodnov in 1947, this type of lune is not squared ([35], p. 62). Therefore, the intermediate of lunes as a minor premiss is very weak, and so it was at the times of Hippocrates of Chios results, and far from being clear and convincing (in fact, now we know that it is false, but it was not false at the Aristotle's times, when it was instead plausible). Indeed, we can easily guess that at the times of Aristotle a great hope of solving the circle-squaring problem using the lunes was widespread, and this fact renders the example acceptable as an illustration of a kind of hypothetical reasoning (abductive!), not necessarily involved in presenting geometrical established results, but just possible results that could be reached thanks to the plausibility of the minor premise (an added middle term)[13] concerning lunes.

[13]Further considerations about the role of additional middle terms in these two Aristotelian examples are illustrated below, Sects. 1.3.1 and 1.3.2.

We need a deeper and better interpretation of Aristotle’s passage. To this aim we need analyze some aspects of Plato’s dialectic, ancient geometrical cognition, and the role of middle terms: by illustrating these aspects, in [27] I tried to convince the reader that we can gain a new positive perspective about the constitutive eco-cognitive character of abduction, just thanks to Aristotle himself. In the present section it was sufficient to stress the eco-cognitive openness indicated by Aristotle with his emphasis on the need in abduction of cognitive externalities—leading away—able to permit reasoners to go beyond that eco-cognitive *immunization* he himself—instead—considered crucial for founding syllogism (see above the last paragraph of Sect. 1).

1.2 Geometry and Logic: The Role of Constructions and Middle Terms in Abduction (“Leading Away”)

Many researchers (for example [9, 19]) contend that Aristotle’s passage above reworks two examples already given by Plato in the *Meno* dialogue [44]. The interpretative conundrum is related to the role played by the middle term: first of all Aristotle points out that abduction is such “when it is clear (δῆλον) that the first term belongs to the middle and unclear (ἄδηλον) that the middle belongs to the third, though nevertheless equally convincing (πιστόν) as the conclusion, or more so”. This situation is not satisfactory because the conclusion/hypothesis is too precarious. Indeed, Aristotle says that to have an abduction an act of *introducing* “something in addition” is necessary, and the addition can also be characterized by more middle terms: “That science is teachable, then, is obvious, but it is unclear whether virtue is a science. If, therefore, BC is equally convincing (πιστόν) as AC, or more so, it is a leading away (ἀπαγωγή) (for it is closer to scientific understanding (ἐγγύτερον γάρ τοῦ ἐπίστασϑαι) because of taking something in addition (τὸ προσειληφὲναι), as we previously did not have scientific understanding (ἐπιστήμή) of AC). Or next, it is leading away ἀπαγωγή if the middle terms between B and C are few (ὀλίγα) [for in this way also it is closer to scientific understanding (εἰδέναι)].”

A more careful analysis of the passage requires a reference to some central Plato’s ideas about dialectic argumentation. Already in the *Meno* dialogue Socrates “dialectically” reflects upon the various relationships between virtue, knowledge, and teachability and also furnishes the example of a geometrical “analysis” (or “method of hypothesis”), so-called in the literature pertaining ancient mathematics.[14]

[14]Cf. for example [18].

1.3 Ἀπαγωγή *and Geometry*

The method of analysis in geometry, already employed by Hippocrates of Chios, can involve, to creatively solve the problem at hand, (1) a *diorism*, which resorts to the finding of the definite conditions under which one construction might be inscribed within another, and (2) a *porism*, which refers to direct or intentional discovery through suitable higher constructions related to the finding of indefinite cases, eventually capable of innumerable solutions, so looking for a higher unifying solution. What is important to note is that in the method of analysis new strategic constructions have to be found: translated in syllogistics terms, this means it is necessary "taking something in addition (τὸ προσειληφὲναι)", as Aristotle says in the passage above, that is a new "middle" (or new "middles").[15]

The activity of finding new geometrical constructions (or new middle terms, Aristotle would say) is clearly a heuristic process[16]—based on a dynamics of subsequent steps—aiming at discovering new geometrical truths, a process which is a case of ἀπαγωγή, that is of *abduction*, also in the modern sense of the word.[17] In a syllogistic perspective, which regards arguments in general, not necessarily geometrical, the method of analysis still resorts to the activity performed for finding the suitable middle term(s) able to substantiate the reasoning at play.

It is absolutely important to note that in Plato the *logico-dialectical* anticipation of the Aristotelian syllogistic relationship between virtue, knowledge, and teachability is directly derived from the geometrical example, as Socrates expressly says in the *Meno* dialogue (see below, the following subsections). From Hippocrates of Chios to Proclus, ἀπαγωγή is the fundamental pre-Euclidean method for solving problems, as a method of discovery, and at the same time also, later on, for proving theorems [19], no surprise that it is implicitly central in Plato and still explicitly present in Aristotle's *Prior Analytics*. In this perspective, we will soon see, the English translations "reduction" and "leading away" both stress the fact that the process involves

[15] Porism is usually translated as lemma or corollary. I am referring here to another meaning that goes deeper into the philosophy of ancient Greek mathematics. In this case porisms are active in solving problems in which it is necessary to adopt new suitable constructions. The most famous collection of porisms of ancient times was the book *The Porisms* of Euclid. This work is lost: the trace survived thanks to the *Collection* of Pappus. Playfair noted that, thanks to porisms, the analysis of all possible particular cases of a proposition would establish that: (1) under some conditions a problem becomes impossible; (2) under some other conditions, indeterminate or related to an infinite number of solutions the problem can be solved. Classical works on porisms are [45, 48]. The concept is controversial and still subjected to studies and interpretations provided by researchers in ancient philosophy: a rich reference to the literature available is given in ([19], pp. 39–40).

[16] I have provided an analysis of heuristics in the light of abductive cognition in [25]. Heuristics, in so far they can be algorithmically rendered, are still rules-based, even if these rules are weaker from the normative point of view, when compared with the logical rules, and typically closer to what actual human reasoners do.

[17] In ([24], chapters two and three), I have illustrated how abductive cognition is also characteristically related to various examples of diagrammatic reasoning (based on porisms, we can say), for example in the case of the discovery of the first non-Euclidean geometries.

a transition from a problem or theorem to another, which, if known or constructed, will make the original problem or theorem evident and solved (or potentially solved).

We have to add that Proclus himself will contend that "Leading away (ἀπαγωγή) is a transition from a problem or theorem to another which, if known or constructed, will make the original problem or theorem evident. For example, to solve the problem of doubling the cube, geometers shifted their inquiry to another on which this depends, namely, the finding of two mean proportionals; and henceforth they devoted their efforts to discovering how to find two means in continuous proportion between two given straight lines. It is reported that the first to effect reduction of difficult (perplexing) geometrical propositions was Hippocrates of Chios who also squared the lune and made many other discoveries in geometry, being a man of genius when it came to constructions, if there ever was one" [Proclus, in Eucl. 213 Translation adapted from [46]; quoted in ([19], p. 23)].

1.3.1 Ἀπαγωγή, Dialectics, and Logic

Socrates and Meno, by constructing the arguments on whether virtue is teachable are engaged in clarifying the following syllogism (obviously valid, but where both the second premiss and the conclusion are far from being reliable or "plausible")

AB Whatever is knowledge, can be taught
BC Virtue (e.g., justice) is knowledge
AC Therefore virtue can be taught.

Faller [10, 11] explains that Socrates, exactly thanks to what Aristotle calls a "leading away" (ἀπαγωγή) argument—that is the introduction of new middles—, had established that since virtue is "good" and "there is nothing good that is not embraced by knowledge, our suspicion that virtue is a kind of knowledge would be well founded" ([44], 87d), consequently, Meno can say: "We must now conclude, I think, that it is; and plainly, Socrates, on our hypothesis that virtue is knowledge, it must be taught" (89c). In the Aristotelian terms exploited in chapter B25:

AB Whatever is knowledge, can be taught
BC Virtue (e.g., justice) is knowledge.

MC Virtue is good[18]
BM Good is knowledge

AC Therefore virtue can be taught

The first premiss is evident, the second uncertain, and the conclusion is even more uncertain. We can arrive—using the Aristotelian words—"closer to scientific understanding (ἐγγύτερον γάρ τοῦ ἐπίστασθαι)", with the introduction of a new

[18] Karasmanis usefully notes that the term "good" is not given in the analogous Aristotelian example I have illustrated in the previous subsection. Aristotle only says that an intermediate term is introduced ([19], p. 37).

term "good" and the propositions "virtue is good" and "good is knowledge" (that is taking something in addition (τὸ προσειληφέναι), which can *possibly* support the second premiss of the original syllogism).

Plato starts from AC, which reflects a situation of ignorance, a hypothesis *to be rendered plausible* or, so to speak, arguable, or even possibly true, instead of its contrary; BC would guarantee the result but it has to be supported. MC, the "leading away" at stake, is the further hypothesis chosen to perform this task: MC results obvious and true.[19]

Then Plato establishes that "good is knowledge" (BM) and concludes that "virtue is knowledge". Through this process (dialectical) AC is rendered—again, using the Aristotelian words —"closer to scientific understanding (ἐγγύτερον γάρ τοῦ ἐπιστασϑαι)": indeed he concludes that virtue can be taught.[20]

As Bedu-Addo observes, Plato clarifies in the *Phaedo* that the additional terms function as further hypotheses that render—in a kind of clear abductive atmosphere, we can say—the main precarious hypothesis "more adequate":

> Cherniss[21] has pointed out, in calling the proposition "virtue is good" a hypothesis, Socrates is exemplifying the "upward path" of the hypothetical method as described in the *Phaedo*: "And when you should have to give account of *the hypothesis itself*, you would do so in the same manner hypothesizing another hypothesis which seemed best of those above, until you came to something adequate (τι ἱχανόν)" (Phaedo, 101d–101e).[22] Thus the proposition "virtue is good" functions as the "something adequate" in the argument "from a hypothesis"; and the entire proof which gives the "account" of the hypothesis "virtue is knowledge'' as recommended at *Phaedo* 101d is to be seen as an instance of "reasoning out the cause", αἰτίας λογισμός, which we are told has to be done if a true opinion is to be converted into knowledge, and which is said to be recollection ([44], 98a), quoted in ([4], p. 10).

Hence, the final reached conclusion that virtue can be taught is characterized as a "true opinion" (ὀρϑή δόξα). Socrates indeed says (we are at the final part of the *Meno* dialogue) "For these [true opinions], so long as they stay with us, are a fine possession, and effect all that is good; but they do not care to stay for long, and run away out of the human soul, and thus are of no great value until one makes them fast with causal reasoning (αἰτίας λογισμός)" ([44], 98a). Actually the syllogisms that are illustrated in these Socratic examples are affected by uncertainty: they indeed lead to a soft cognitive commitment to the conclusion/hypothesis because this is a "true opinion" (ὀρϑή δόξα). Plato/Socrates underlines the potential withdrawability of true opinions and so their provisionality.

---

[19]This proposition corresponds to that *arche* (ἀρχή) which was so called, in the case of the geometrical analysis, by Hippocrates of Chios [cf. ([23], chapter four)].

[20]In sum, to render more acceptable and justifiable that virtue is teachable it has been necessary to analyze its nature: "what is virtue"; it has been necessary the method of hypothesis to examine the features of an obscure subject.

[21]Cherniss ([7], p. 140).

[22]The following is another translation that adopts the term principle instead of hypothesis (for ὑπόϑεσις), but that expresses the same argument: "And when you had to give an explanation of the principle, you would give it in the same way by assuming some other principle which seemed to you the best of the higher ones, and so on until you reached one which was adequate " ([43],101d–101e).

As I have anticipated above in Sect. 1 I think Peirce is crystal clear about the issue, presenting this case, adopted by Aristotle, as a typical way of considering abduction. He says: “He [Aristotle] immediately proceeds to give the needed examples. The first is this: ‘Now that comprehension [that is, knowledge] is capable of being taught is plain; but that virtue is comprehension is not known. If, however, this is as antecedently likely or more so, than that virtue should be capable of being taught (which, it seems needless to say, everybody knows to be the fact), then there is ground for the abduction; since we are brought by the hypothesis ‘τὸ προσειλήφἐναι’[23] nearer to a comprehension of virtue being capable of being taught, than we were before’. This seems very clear. […] He now gives another example to illustrate the case in which the hypothesis brings us nearer to comprehension because, to use his phrase, ‘the middles are few’, that is, it seems near to first principles” ([39], 7.250). Peirce concludes “This interpretation is that abduction is nothing but an ordinary syllogism of the first figure, when we are not sure of the minor premiss, but still are more inclined to admit it than we should be to admit the conclusion if the latter were not a necessary consequence of the former” ([39], 7.251).[24]

Let us come back to the *Meno* dialogue. Meno accepts the final reached conclusion (AC, see above in this subsection) that virtue can be taught but Socrates is not satisfied: to solve the problem we need—still in Aristotelian words—another “leading away”. Indeed Socrates initiates a second argument consisting of a further hypothesis, that “if virtue is teachable, then there would be teachers of it”. Unfortunately, because of the empirical fact that there are no teachers of virtue, virtue is not teachable, a conclusion which conflicts with the previous one about teachability.

First of all we have to note and remember that (1) in the example about virtue Plato adopts exactly the same method used in geometrical “analysis”. We also have to stress that (2) we reached two conflicting conclusions (already available, one statement and its negation) and further steps would have to be performed to execute the *cutdown* process (cf. [29], Chapter one, Sect. 1.2), that is the activity of “pruning” hypotheses, still to arrive “closer to scientific understanding (πάντως γάρ ἐγγύτερον εἶναι σύμβαίνει τῆς ἐπιστήμης)”, that is to a unique conclusion (the best result, which echoes abduction as the best explanation). Exactly in

[23] We have already seen that this expression means “taking something in addition”; see above in Sect. 1 the English translation of chapter B25 that contains this expression.

[24] This first example given by Aristotle and derived from Plato is also unfavorably commented by Peirce, and I think the negative verdict does not concern its formal congruity with the idea of abduction expressed in chapter B25, but instead the philosophical content and quality: “But when we come to the examples, the ordinary interpretation reduces the latter, at least, to nonsense. The first becomes, Comprehension can be taught, Virtue is comprehension; ∴ Virtue can be taught. In the first place, this is a *petitio principii*, or very near to one since there is no way of proving that virtue is comprehension, except by its being taught. In the next place, few in Aristotle’s time had used this absurd argument; it had scarcely been seriously doubted, what all experience shows, that virtue can be taught. A very few ethical writers of modern times have denied it; but it had hardly been denied then, except as a temporary shift in debate. A philosopher who, like Socrates, maintained that it was better to do wrong, knowing it, could not doubt that righteousness could be taught” ([39], 7.251).

the spirit of Peircean original perspective on abduction, we have to select (and so to prefer) one of the two conflicting conclusions.

Again, let me stress that Plato's argumentation about virtue is the dialectic analogue of a diorismic/poristic geometrical process, which in turn substantiates the Aristotelian "taking something in addition" (τὸ προσειλῆφέναι), where various strategies can be further activated: various kinds of arguments (for example the reaching of evident higher hypotheses from which the initial one can be deduced), considerations of simplicity, looking for consequences (for example in terms of empirical ascertainments and testing), which are able not only to create new cognitive perspectives (fill-up aspect) but also to select (cutdown aspect) the multiple or conflicting flow of results.

1.3.2 Geometry and Logic Intertwined: Ἀπαγωγή and Its Eco-Cognitive Openness

Geometrical analysis initially transforms a given problem into one that is more abstract and general: even if there are conflicting views in the available literature on the subject,[25] we can say that diorisms and porisms (often consisting in the depicting of locus[26] problems) favor a form of further geometrical cognition devoted—thanks to the study of auxiliary objects—to finding the conditions of possibility of an actual process of subsequent diagrammatic constructions, in turn finalized to solve the problem. Diorisms aim at determining the overall properties of the solutions, and so represent a wide range of mathematical activities, which "lead away" from the problem at hand to other unexplored porismic territories (diagrams and sentential proofs for example, but, for the sake of generality of various cognitive processes, we can also add other model-based or manipulatory activities totally eco-cognitively open). In the diorismic/poristic stage, the geometrician exploits the adopted auxiliary objects to show that a single solution is always possible, or if not, the limitations of the process or how many solutions there may be and how they are arranged [47].

The process performs a *reduction* of the problem— caused by the hypothetical question to be solved—to another one (again, it is the Aristotelian syllogistic "leading away"), which we expect will enable us to solve the original problem (I have already said that in the Aristotelian passage above the word ἀπαγωγή is often translated with "reduction", and that we need interpret reduction as the transition to another cognitive sub-process and not as the *reduction ad absurdum*).

In the case of Plato's second problem—the geometrical one—we are to determine whether a certain rectilinear figure could be constructed along the diameter of a circle examining it by means of "a certain helpful hypothesis" (that is by means of an additional term, in Aristotelian words), expressed in the following passage (see the emphasis I have added), where a state of ignorance is immediately declared". [I have

[25]Cf. above, footnote 15 at p. 88.

[26]It is interesting to note that the term *topoi* (in Latin *loci*) migrates to Aristotle's rhetoric and later rhetoricians' studies, probably parasitic of its origin in geometrical analysis [11].

already stressed in the previous subsection that it is important to note that in Plato it is just the “clarification” of the dialectic relationship between virtue, knowledge, and teachability, which is directly derived from the geometrical example, as Socrates expressly says]:

> So it seems we are to consider what sort of thing it is of which we do not yet know what it is! Well, the least you can do is to relax just a little of your authority, and allow the question – whether virtue comes by teaching or some other way – to be examined by means of hypothesis. I mean by hypothesis what the geometricians often do in dealing with a question put to them; for example, (86e) whether a certain area is capable of being inscribed as a triangular space in a given circle: they reply – “*I cannot yet tell whether it has that capability; but I think, if I may put it so, that I have a certain helpful hypothesis for the problem, and it is as follows: If this area is such that when you apply it to the given line [as a rectangle of equal area] of the circle you find it falls short by a space similar to that which you have just applied, then I take it you have one consequence, and if it is impossible for it to fall so, then some other. Accordingly I wish to put a hypothesis, before I state our conclusion as regards inscribing this figure in the circle by saying whether it is impossible or not*” ([44], 86e, 87a).

Let me reiterate that Socrates explicitly analogizes his reasoning about virtue to the one used in the geometrical example, and we can reasonably guess that the source of the analogy is exactly the just illustrated geometrical example:[27]

> In the same way with regard to our question about virtue, since we do not know either what it is or what kind of thing it may be, we had best make use of a hypothesis in considering whether it can be taught or not, as thus: what kind of thing must virtue be in the class of mental properties, so as to be teachable or not? In the first place, if it is something dissimilar or similar to knowledge, is it taught or not – or, as we were saying just now, remembered? (cit., 87b).

To determine whether a certain rectilinear figure could be constructed along the diameter of a circle Socrates establishes the hypothesis which I have emphasized in the first passage above from the *Meno* dialogue: the hypothesis needs be worked thanks to a diagrammatic process, a “leading away”, which opens up the reasoning to an eco-cognitive dimension, which in our case corresponds to the ἀπαγωγή: an abduction, endowed with its degrees of uncertainty. The echo of this reference to the importance of diagrams in analyzing reasoning is still vivid in Peirce: “I said, Abduction, or the suggestion of an explanatory theory, is inference through an Icon” ([38], p. 276).

A brief note on recent cognitive rich research on diagrammatic geometrical reasoning has to be introduced. Fresh studies have shown that false premisses [also due to the presence in models/diagrams of both substantive and auxiliary assumptions, indeed spurious problematic sub-diagrams and new “individuals” can pop-up

[27] A strict relationship between geometry and dialectics stills echoes in Proclus: “[…] mathematics reaches some of its results with analysis, others by synthesis, expounds some matters by division, others by definition, and some of its discoveries binds fast by demonstration, adapting these methods to its subjects and employing each of them for gaining insight into mediating ideas. Thus its analyses are under the control of dialectic, and its definitions, divisions, and demonstrations are of the same family and unfold in conformity with the way of mathematical understanding. It is reasonable, then, to say that dialectic is the capstone of the mathematical sciences” ([46], 43, p. 35).

at any step of geometric constructions ([8], p. 105)] are not exploited in the cognitive abductive process, because, in the various heuristics, only the *co-exact* properties are exploited. As I have described in Sect. 4.4 of ([29], chapter four) the notion of co-exact properties, introduced by [32], is worth to be further studied in fields that go beyond the realm of deductive processes of classical geometry, in which it has been nicely underscored, so usefully touching various discovery cognitive processes.[28] [34] illustrates that in Euclid's deductive framework diagrams contribute to proofs only through their co-exact properties: I suggest that this is also typical of diorismic/porismic processes and of their creative counterparts, exactly endowed with an objection-refuting role.

In the Aristotelian (and Platonic) perspective (see Chap. B25 of the *Prior Analytics*) I have delineated in this section we can definitely conclude that the general concept of abduction must be seen as constitutively and widely *eco-cognitive-based*. Indeed, by contrast, we have to remember that Aristotle says, in the passage I have already quoted and that I am reporting again, that a valid syllogism—by necessity—is instead not at all open to something "external": "A deduction (σύλλογισμòζ) is a discourse (λόγοζ) in which, certain things having been supposed, something different from the things supposed results of necessity because these things are so. By 'because these things are so'. I mean 'resulting through them,' and by 'resulting through them' I mean '*needing no further term from outside* (ἔξωϑεν) *in order for the necessity to come about*' " ([1], A1 24, 20–25, p. 2) (emphasis added).

Even if in this article I cannot illustrate in detail the diagrammatic constructions, which make possible to afford the geometrical problem illustrated by Plato-Socrates,[29] it is clear that, in syllogistic terms, the geometrical diagrammatic process, as well as the analogue argumentation about virtue, are ways of finding a "middle" ground that solves the problems at hand. Aristotle concludes, in *Posterior Analytics* "Thus it results that in all our searches we seek either if there is a middle term or what the middle term is. For the middle term is the explanation, and in all cases it is the explanation which is being sought" ([2], B, 90a, 5, p. 48).

At this point there is clear evidence that both Socrates' examples are recalled, with slight differences, in Aristotle's celebrated passage about abduction from chapter B25 of *Prior Analytics*.

Let us come back to the geometrically puzzling example present in the Aristotelian passage, already reported above, involving the effort to square the circle through the lunes, a problem typical, together with the one related to the reduction of the famous Delian problem, of the geometrical research deriving from Hippocrates of Chios:

DE Whatever is rectilinear, can be squared

[28] Manders' definition describes the co-exact properties "as those conditions unaffected by some range of every continuous variation of the diagram" and the exact ones as "those which, for at least some continuous variation of the diagram, obtain only in isolated cases" [32]. "Diagrams of a single triangle, for instance, vary with respect to their exact properties. That is, the lengths of the sides, the size of the angles, the area enclosed, vary. Yet with respect to their co-exact properties the diagrams are all the same. Each consists of three bounded linear regions, which together define an area" ([34], p. 264).

[29] An interesting reconstruction is given in [10].

EF A circle can be transformed into a rectilinear figure by the intermediate of lunes[30]
DF Therefore, a circle can be squared
D = square, E = rectilinear figure, F = circle

The first premiss is known and true, the second is uncertain, the conclusion even more uncertain: a “leading away”, towards the lunes, has to start. In the above syllogism we introduce a new term (N = lune) and two new premisses “EN = the lunes become rectilinear” and “NF = the circle is a sum of lunes”: *thanks to and together with* the related diagrammatic constructions, the schema becomes

DE Whatever is rectilinear, can be squared
EF A circle can be transformed into a rectilinear figure by the intermediate of lunes

---

EN The lunes become rectilinear
NF The circle is a sum of lunes

DF Therefore, a circle can be squared.

The new additional premiss, fruit of a “leading away”, aims at supporting the second uncertain premiss EF to approximate to knowledge and so to solve the problem. “Our problem, which is squaring the circle, is reduced to that of squaring the lunes, John Philoponus sees in this way the Aristotelian argument by *apagoge*. […] I think it is now plain that Aristotle’s *apagoge* (apart from its syllogistic form) fits very well with Proclus’ *apagoge*” ([19], p. 27). Moreover, Aristotle clearly notes, we have to deal with one or few new intermediate terms: the importance of minimality of abductive cognition is prefigured. The concept of abduction is finally established: it is leading away (ἀπαγωγή), Aristotle concludes, if the middle terms between B and C are few (ὀλίγα) [for in this way we are also closer to scientific understanding εἰδέναι]. The efficacy of the abductive procedure is thus dependent on a minimum of middle terms, because too many moves will generate excessive distance for the argument to be convincing. In case of multiple kinds of middle additional terms which lead to different conclusions we still have to select/discriminate both the appropriate/productive additional middles and the best related final result (cutdown problem).

Also the above mathematical second example furnished by Aristotle is negatively considered by Peirce, but also in this case I think the bad judgement does not regard its coherence with the idea of abduction expressed in chapter B25, but instead the mathematical content and quality, considered unsatisfactory:

> The other example is still worse. It becomes, Whatever is equal to a rectilinear figure can be squared, Every circle is equal to a rectilinear figure; ∴ Every circle can be squared. We here naturally understand by “equal to a rectilinear figure”, equal to a rectilinear figure, constructible or inconstructible. But in that case, the minor premiss, instead of being not known, is the most evident thing in the world; while the major premiss which ought to be

---

[30]See again the considerations about the classic problem of squaring the circle with the help of the lunes given at p. 86.

> manifest, is far from being so; for if a figure cannot be constructed it cannot be squared. Supposing however that by a rectilinear figure is meant one that can be constructed, which must have been the meaning, since Aristotle says that it is almost known through lunes, who ever used such a ridiculous argument? And how can Aristotle say, as he does, that lunes in any way help the matter, or are at all relevant? Whatever bearing lunes were supposed to have upon the quadrature of the circle disappears entirely from this representation. Nothing can be more utterly unlike Aristotle's usual examples, which bring up in vivid aptness actual reasonings well known to his scholars ([39], 7.252).

Perplexities are shown about the major premiss, considered questionable, perhaps based on the fact that if a figure cannot be somewhat constructed, it cannot be squared even if it is rectilinear. Perplexities are shown about the minor premiss that may be Peirce considers entirely evident because new methods can permit to square the circle beyond the use of compass and straightedge. Not only, Peirce adds the following sentence: "Mathematics was not Aristotle's strong point, and possibly he did not clearly understand that it was only two or three special lunes that Hippocrates had squared". In reality, these Peircean comments, which in my opinion manifest, I repeat, a bad judgement that concerns the mathematical content and quality of the example, but not its coherence with the idea of abduction expressed in chapter B25, do not consider that, as I have already said in Sect. 1.1 above, at the times of Aristotle a great hope of solving the circle-squaring problem using the lunes was surely widespread. Hence, the example is reasonably acceptable as an illustration of a kind of hypothetical abductive cognition, obviously not involved in showing geometrical established results, but just a "possible" result that could be reached thanks to the increased plausibility of the minor premise, thanks to the added middle term concerning lunes.

1.4 Dialectics, Rules of Interrogation, Syllogisms: Dialectical Logic Versus Syllogistic Logic?

We know that in the light of classical logic abduction is the fallacy of the affirming the consequent. As I have already illustrated in Sect. 1, in the case of the Aristotelian chapter B25, abduction does not have to be discussed keeping in mind this fallacious schema. What is at stake in chapter B25 of *Prior Analytics*, in which Peirce envisages the first appearance of abduction, is the double face of abduction that can be considered either (1) the classification of a certain "unclear" dynamic argument in a *context-free* sequence of three propositions; or (2) the introduction in a similar "unclear" dynamic three-propositions argument (in this case no longer *context-free*) of few new middle terms. Indeed, chapter B25 is built taking advantage of the examples provided by Plato. These examples are usually seen in a "dialectic" perspective, where both propositional and geometrical aspects are illustrated. It is important to note that Aristotle seems to sterilize the dialectic background pointing to the general and abstract role of those "propositional" externalities which are represented by the additional middle terms that substantiate the process of "leading away".

To try to clarify the above difficulties we can report a dispute about the status of the Aristotelian view of logic and fallacies, which might also acquire some further light thanks to the study illustrated in the present article. The dispute is due to Hintikka and Woods. Hintikka [16] contends that in the *Topics* and *On Sophistical Refutations* the aim was practical: “Aristotle did not only want to study knowledge-seeking interrogative games for abstract theoretical purposes. He wanted to show how to win in such games” ([17], p. 242),[31] and “Aristotle is still in the two *Analytics* thinking of logical inferences as steps in a questioning process” (cit., p. 243). I think that this perspective on logical inferences in terms of interrogative games favors a reading of Aristotle under the dialectic lens of Plato and overlooks the syllogistic spirit which pervades the study of arguments: in sum, Hintikka thinks that Aristotle treats the whole inquiry as an interrogative process and is looking for a theory of interrogative dialectical reasoning intended as a theory of ampliative reasoning, where what we now call deductive routines are just part and parcel of the dialectical/interrogative process. Following Hintikka, fallacies (for instance begging the question and many questions) are just considered by Aristotle violations of rules of interrogation.

It is interesting to note, by the way, that Hintikka’s early (indeed foundational) attachment to a game-theoretic approach to logic involved the idea that dialectical interrogation rules are essential to the definition of logical particles (“all”, “some”) and logical relations (entailment, proof). Some scholars attribute this view to Aristotle [cf. [33] and ([52], chapter three)]. The big difference between Hintikka and Woods’ approach seems to resort to the fact that Woods sees Aristotle as presenting two logics—one for syllogisms-as-such, in which nothing dialectical occurs, and another for syllogisms-in-use, in which dialectical considerations are sometimes vitally involved. Hintikka, on the other hand, thinks that dialectical factors are of central concern throughout.

Indeed, Woods [50] favors a reading of Aristotle, which attributes proper value to his invention of syllogism, contending that he clearly distinguished syllogism from dialectics: “Arguments in the narrow sense [not dialectical] stand starkly apart. They are not social events. They are not events of any kind. They are finite sequences of linguistic objects which Aristotle calls propositions. When they meet certain conditions, they are *syllogisms*”, so “there is no great harm in distinguishing between Aristotle’s *dialectical logic* and his *syllogistic logic*. But we should not lose sight of the point that these are disjoint conceptions of logic”: the logic of syllogism is the theoretical core of a successful dialectic in the broad sense, but we need to establish the distinction. In sum Woods objects that Aristotle develops both a theory of two-person argument (that is a theory about rules of interrogation in Hintikka’s sense, where an interrogative dialectical argument is a series of alternating speech acts between opponents, questioner and respondent), *and* a theory of syllogism, intended as a context free sequence of three categorical propositions. Woods concludes “Certainly Aristotle would allow that these [begging the question and many questions]

[31]This Hintikka’s article “What was Aristotle doing in his early logic, anyway? A reply to Woods and Hansen”, belongs to the published intellectual fight between Hintikka and Woods and it is an answer to Woods and Hansen’s “Hintikka on Aristotle’s Fallacies” [55].

are interrogative rule-violations, but it is not this that makes them fallacies" ([51], pp. 497–498). I think that this perspective helps us to read Aristotle's chapter B25 on ἀπαγωγή as a chapter which certainly reverberates some aspects of dialectics but in the sense of syllogistic logic, where abduction is just characterized by the breaking of high degrees of certainty. In sum, Aristotle takes pains to strip away all strict dialectical considerations.[32]

Leaving aside the interesting *querelle* about the status of fallacies in Aristotle, we can nevertheless say that the study of Aristotelian ἀπαγωγή illustrated above provides support to Woods' conviction about a theory of syllogism "as a context free sequence of three categorial propositions": indeed it clearly results that in valid syllogisms there is no room for broad eco-cognitive acts of "leading away", such as the ones that are instead illustrated in chapter B25 of the *Prior Analytics*. At this point there is clear evidence that both Socrates' examples are recalled, with slight differences, in Aristotle's celebrated passage about abduction from the *Prior Analytics*, but they are embedded in a very different theoretical framework, in which the classification of certain "unclear" dynamic arguments in both *context-free* and *non context-free* sequences of three propositions dominates.

2 Abduction and the Optimization of Eco-Cognitive Situatedness: Discoverability and Diagnosticability Explained

When I say that abduction can be knowledge-enhancing[33] I am referring to various types of new produced knowledge of various novelty level, in absence of an empirical evaluation phase, or inductive phase, as Peirce called it. Some cases of new knowledge produced in science (for example, conventions and intermediate models used in research settings), are cases of knowledge enhancing abduction. However, also knowledge produced in an artificial game thanks to a smart application of strategies or to the invention of new strategies and/or heuristics has to be seen as the fruit of knowledge enhancing abduction.

[32]Woods further observes that the great achievement of *Prior Analytics* is metalogical. It provides an almost perfect (and easily repairable) proof of the semi-decidability of validity for arguments framed under syllogistic constraints. The proof rules for this are both syllogistic and "common", and there is nothing remotely dialectical about the procedure in the Hintikkean sense.

[33]This means that abduction is not necessarily ignorance-preserving (reached hypotheses would always be "presumptive" and to be accepted they always need empirical confirmation), as contended by Gabbay and Woods (see [51]). Abduction can creatively build new knowledge by itself (that as an inference not necessarily characterized by an empirical evaluation phase, or inductive phase), as various examples coming from the area of history of science and other fields of human cognition clearly show. I better supported my claim about the knowledge enhancing character of abduction in the recent [27, 28]. On this issue see also above footnote 2. Woods has recently enriched, modified, and moderated his views of ignorance-preservation, see [53].

I contend that to reach selective or creative good abductive results efficient strategies have to be exploited, but it is also necessary to count on an environment characterized by what I have called *optimization of eco-cognitive situatedness*, in which that eco-cognitive openness already envisaged by Aristotle thanks to the emphasis on "leading away"[34] is fundamental [28]. To favor good creative and selective abduction reasoning strategies must not be "locked" in an external restricted eco-cognitive environment such as in a scenario characterized by fixed definitory rules and finite material aspects (an artificial game, Go or Chess, for example), which would function as cognitive mediators able to constrain agents' reasoning.[35] In brief, the optimization of eco-cognitive situatedness concerns the substantial problem of *discoverability* and *diagnosticability*, almost totally disregarded in the literature on abduction (and just sketched by Peirce himself).[36]

Research on abduction has frequently emphasized the fruitful role of cognitive openness. Hendricks [15] consider trans-paradigmatic abduction a form of discovery in which a guessed hypothesis transcends the prompt empirical agreement between two paradigms. The paradigms are presumed to belong to the same field (for example physics) where one of the fields is well established and the other is emerging (for example classical and quantum physics):

> A case in point would be the formulation of the hypothesis of electron spin. Bohr considered the spin conjecture as a welcome supplement to the current magnetic core theory. Pauli remained rather skeptical pertaining to the spin hypothesis due to the fact that it actually required the theory of quantum mechanics for its proper justification, which was not part of the background knowledge at the time of the conjecture. In such cases two paradigms are competing and the abduction is then dependent upon whether the conjecture is made within the paradigm or outside it. Hence we distinguish between paradigmatic and trans-paradigmatic abduction (cit., p. 287).

Furthermore, people draw on different domains of knowledge to arrive to an abductive conclusion thanks to what [14] call "transepistemic abduction" (TeA), which illustrates how two agents, in order to successfully explain a phenomenon, reason across two very distant cognitive fields (for example computational and psychosocial domains) despite each agent being ignorant of the other domain knowledge. The authors themselves acknowledge that TeA represents a case that is partially concerned with my eco-cognitive perspective: "TeA may not necessarily accommodate wider understandings of abduction like the eco-cognitive model proposed by Magnani. For example, TeA may not necessarily encompass perceptions, aesthetic decisions or moral judgements in the way that a eco-cognitive view of abduction might".

Another interesting procedure that can refer to higher abductive processes in need of cognitive openness is the chunk-and-permeate method [6], in which consideration is given to conditions under which mutually incompatible well-grounded theories can interact to bring forth solutions to problems which neither theory can solve

[34]Cf. the previous section of this article.

[35]More details concerning the role of locked and unlocked strategies are illustrated in the recent [30].

[36]I plan to devote part of my future research to study these aspects of abductive cognition.

on its own. This method introduces a paraconsistent reasoning strategy, in which information is broken up into chunks, and a limited amount of information is allowed to flow between chunks, and it is applied to model the reasoning employed in the original infinitesimal calculus.

It is now useful to provide a short introduction to the concept of eco-cognitive openness from a logical point of view. The new perspective inaugurated by the so-called *naturalization of logic*[37] contends that the normative authority claimed by formal models of ideal reasoners to regulate human practice on the ground is, to date, unfounded. It is necessary to propose a "naturalization" of the logic of human inference. Woods holds a naturalized logic to an adequacy condition of "empirical sensitivity" [51]. A naturalized logic is open to study many ways of reasoning that are typical of actual human knowers, such as for example fallacies, which, even if not truth preserving inferences, nonetheless can provide truths and productive results. Of course one of the best examples is the logic of abduction, where the naturalization of the well-known fallacy "affirming the consequent" is at play. Gabbay ([13], p. 81) clearly maintain that Peirce's abduction, depicted as both (a) a surrender to an idea, and (b) a method for testing its consequences, perfectly resembles central aspects of practical reasoning but also of creative scientific reasoning.

It is useful to refer to my recent research on abduction [28], which stresses the importance in good abductive cognition of the already quoted *optimization of situatedness*: abductive cognition is for example very important in scientific reasoning because it refers to that activity of creative hypothesis generation which characterizes one of the more valued aspects of rational knowledge. The study above teaches us that situatedness is related to the so-called eco-cognitive aspects, referred to various contexts in which knowledge is "traveling": to favor the solution of an inferential problem—especially in science but also in the case of other abductive problems, such as diagnosis—the richness of the flux of information has to be maximized.

It is interesting to further illustrate this problem of optimization of eco-cognitive situatedness taking advantage of simple logical considerations. Let $\Theta = \{\Gamma_1, \ldots, \Gamma_m\}$ be a theory, $P = \{\Delta_1, \ldots, \Delta_n\}$ a set of true sentences corresponding—for example—to phenomena to be explained and $\Vdash$ a consequence relation, usually—but not necessarily—the classical one. In this perspective an abductive problem concerns the finding of a suitable improvement of $A_1, \ldots, A_k$ such that $\Gamma_1, \ldots, \Gamma_m, A_1, \ldots, A_k \Vdash_L \Delta_1, \ldots, \Delta_n$ is *L-valid*. It is obvious that an improvement of the inputs can be reached both by additions of new inputs but also by the modification of inputs already available in the given inferential problem. I contend that to get good abductions, such as for example the creative ones that are typical of scientific innovation, the input and output of the formula

$$\Lambda_1, \ldots, \Lambda_i, ?_I \Vdash_L^X \Upsilon_1, \ldots, .\Upsilon_j \tag{1}$$

(in which $\Vdash_L^X$ indicates that inputs and outputs do not stand each other in an expected relation and that the modification of the inputs $?_I$ can provide the solution) have to be

[37] I have illustrated this new project in [26]. See also above the last part of Sect. 1.

thought as *optimally positioned*. Not only, this optimality—for example in scientific hypothetical reasoning—is made possible by a *maximization of changeability* of both input and output; again, not only inputs have to be enriched with the possible solution but, to do that, other inputs have usually to be changed and/or modified.[38]

Indeed, in our eco-cognitive perspective, an "inferential problem" can be enriched by the appearance of new outputs to be accounted for and the inferential process has to restart. This is exactly the case of abduction and the cycle of reasoning reflects the well-known nonmonotonic character of abductive reasoning. Abductive consequence is ruptured by new and newly disclosed information, and so defeasible. In this perspective abductive inference is *not only* the result of the modification of the inputs, but, in general, actually involves the intertwined modification of both input and outputs. Consequently, abductive inferential processes are highly *information-sensitive*, that is the flux of information which interferes with them is continuous and systematically human(or machine)-promoted and enhanced when needed. This is not true of traditional inferential settings, for example proofs in classical logic, in which the modifications of the inputs are *minimized*, proofs are usually taken with "given" inputs, and the burden of proofs is dominant and charged on rules of inferences, and on the smart choice of them together with the choice of their appropriate sequentiality. This changeability first of all refers to a wide psychological/epistemological openness in which knowledge transfer has to be maximized.

In sum, considering an abductive "inferential problem" as symbolized in the above formula, a suitably anthropomorphized logic of abduction has to take into account a continuous flux of information from the eco-cognitive environment and so the constant modification of both inputs and outputs on the basis of both

1. the *new information available*,
2. the *new information inferentially generated*, for example new inferentially generated inputs aiming at solving the inferential problem.

To conclude, optimization of situatedness is the main general property of logical abductive inference, which—from a general perspective—defeats the other properties such as minimality, consistency, relevance, plausibility, etc. These are special subcases of optimization, which characterize the kind of situatedness required, at least at the level of the appropriate abductive inference to generate the new inputs of the above formula.

In science the optimization of eco-cognitive openness and situatedness surely also concerns the maximization of the information that has to made available, which epistemologically characterizes the tradition of modern science. This fact favors a new attention to the current fragile status human abductive cognition in scientific enterprise. Is the current situation of scientific research facing with problems that jeopardize the optimization of situatedness needed by good abductive cognition? In chapter eight of my book [29] I have provided a first illustration of some negative aspects that have to be stressed when dealing with the current challenges to the creative productiveness of human abductive cognition in science. Indeed, I have

[38] More details are illustrated in ([28], section three).

introduced the hot problem of the contemporary emergence of disparate kinds of "epistemic irresponsibility", related to the commodification and commercialization of science, the marketing of technoscientific products, the impoverishment of the so-called "epistemological niches", which show that human fruitful abductive cognition in science is increasingly assaulted and jeopardized, and at the same time human creativity possibly endangered.

3 Conclusion

I this article I have centered the attention on two themes which are appropriate to stress, in an interdisciplinary perspective and taking advantage of my EC-model of abduction, the philosophical, logical, cognitive, and pragmatic aspects of building hypotheses thanks to abductive cognition. First of all I illustrated a new interpretation of Aristotle's seminal work on abduction and stressed his emphasis on the need of a related situation of eco-cognitive openness, beyond that kind of eco-cognitive immunization against external cognitive interferences Aristotle himself considered necessary in valid syllogisms. Moreover, by enriching the concepts I have introduced in the first section, I illustrated the concept of *optimization of the eco-cognitive situatedness*, describing it as one of the main characters of the abductive inferences to new hypotheses, also stressing the specific case of science. In the background I also recommended that thanks to the two subjects I have illustrated we can gain a new positive perspective about the "constitutive" eco-cognitive character of abduction and its intellectual usefulness in providing a first insight into the fundamental role of "discoverability" and "diagnosticability" in abductive successful cognitive results.

Acknowledgements Parts of this article are excerpted from chapters five and six of the book *The Abductive Structure of Scientific Creativity. An Essay on the Ecology of Cognition*, Springer, Cham, 2017. For the instructive criticisms and precedent discussions and correspondence that helped me to develop my analysis of abductive cognition I am indebted and grateful to John Woods, Atocha Aliseda, Woosuk Park, Luís Moniz Pereira, Paul Thagard, Ping Li, Athanassios Raftopoulos, Michael Hoffmann, Ahti-Veikko Pietarinen, Gerhard Schurz, Walter Carnielli, Akinori Abe, Yukio Ohsawa, Cameron Shelley, Shahid Rahman, Cristina Barés Gómez, Matthieu Fontaine, Gerhard Minnameier, Pat Langley, Oliver Ray, John Josephson, Ferdinand D. Rivera and to my collaborators Selene Arfini and Riccardo Dossena.

References

1. Aristotle: Prior Analytics. Translated by R. Smith. Hackett Publishing Company, Indianapolis/Cambridge (1989)
2. Aristotle. Posterior Analytics, 2nd edn, translated by J. Barnes. Clarendon Press,Oxford (1993)
3. Barés-Gómez, C., Fontaine, M.: Between sentential and model-based abductions. A dialogical approach. Logic J. IGPL **27**(6), 777–811 (2020). Special Issue on "Formal Representations in

Model-Based Reasoning and Abduction", Nepomuceno, A., Magnani, L., Salguero, F., Barés-Gómez, C., Fontaine, M. (eds.) Online first: https://doi.org/10.1093/jigpal/jzz033
4. Bedu-Addo, J.T.: The role of the hypothetical method in the "Phaedo". Phronesis **24**(2), 111–132 (1979)
5. Bellucci, F.: Abduction in Aristotle. In: Gabbay, D., Magnani, L., Park, W., Pietarinen, A. (eds.) Natural Arguments. A Tribute to John Woods, pp. 551–564. College Publications, London (2019)
6. Brown, B., Priest, G.: Chunk and permeate, a paraconsistent inference strategy. Part I: the infinitesimal calculus. J. Philos. Logic **33**(4), 379–388 (2004)
7. Cherniss, H.: Some war-time publications concerning Plato. I. Am. J. Philology **68**(2), 113–146 (1947)
8. Crippa, F.: To prove the evident. On the inferential role of Euclidean diagrams. Theor. Sci. **31**(2), 101–112 (2009)
9. Faller, M.: Plato's Philosophical Use of Mathematical Analysis. Doctoral dissertation, University of Georgia, Athens, GA, USA (2000)
10. Faller, M.: Plato's geometrical logic. Proc. Soc. Ancient Greek Philos. Online (2003a)
11. Faller,M.: The Origin of Peirce's Abduction in Plato's Analytic Method. Manuscript, Online (2003b)
12. Flórez, J.A.: Peirce's theory of the origin of abduction in Aristotle. Trans. Charles S. Peirce Soc. **50**(2), 265–280 (2014)
13. Gabbay, D.M., Woods, J.: The Reach of Abduction. North-Holland, Amsterdam (2005)
14. Gibson, A., Bruza, P.: Transepistemic abduction: reasoning across epistemic domains. Logic J. IGPL. Special Issue on "Formal Representations in Model-Based Reasoning and Abduction", Nepomuceno, A., Magnani, L., Salguero, F., Barés-Gómez, Fontaine, M. (eds.). Online first: https://doi.org/10.1093/jigpal/jzz035
15. Hendricks, F.V., Faye, J.: Abducting explanation. In: Magnani, L., Nersessian, N.J., Thagard, P. (eds.) Model-Based Reasoning in Scientific Discovery, pp. 271–294. Kluwer Academic/Plenum Publishers, New York (1999)
16. Hintikka, J.: The fallacy of fallacies. Argumentation **1**, 211–238 (1987)
17. Hintikka, J.: What was Aristotle doing in his early logic, anyway? A reply to Woods and Hansen. Synthese **113**, 241–249 (1997)
18. Hintikka, J., Remes, U.: The Method of Analysis. Its Geometrical Origin and Its General Significance. Reidel, Dordrecht (1974)
19. Karasmanis, V.: Ἀπαγωγή: Hyppocrates of Chios and Plato's hypothetical method of the Meno. In: Longo, A., Forno, D.D. (eds.) Arguments from Hypotheses in Ancient Philosophy, pp. 21–42. Bibliopolis, Naples (2011)
20. Kraus,M.:Charles S. Peirce theory of abduction and the Aristotelian enthymeme from signs. In: van Eemeren, F.H., Blair, J.A., Willard, C., Henkemans, A.F.S. (eds.) Anyone Who Has a View. Theoretical Contributions to the Study of Argumentation, pp. 237–254. Kluwer Academic Publishers, Dordrecht/Boston/London (2003)
21. Longo, A., Forno, D.D. (eds.): Arguments from Hypotheses in Ancient Philosophy. Bibliopolis, Naples (2011)
22. Magnani, L.: Abduction, Reason, and Science. Processes of Discovery and Explanation. Kluwer Academic/Plenum Publishers, New York (2001a)
23. Magnani, L.: Philosophy and Geometry. Theoretical and Historical Issues. Kluwer Academic Publisher, Dordrecht (2001b)
24. Magnani, L.: Abductive Cognition. The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning. Springer, Heidelberg/Berlin (2009)
25. Magnani, L.: Are heuristics knowledge-enhancing? Abduction, models, and fictions in science. In: Ippoliti, E. (ed.) Heuristic Reasoning, pp. 29–56. Springer, Heidelberg/Berlin (2014)
26. Magnani, L.: Naturalizing logic. Errors of reasoning vindicated: logic reapproaches cognitive science. J. Appl. Logic **13**, 13–36 (2015a)
27. Magnani, L: The eco-cognitive model of abduction. Ἀπαγωγή now: Naturalizing the logic of abduction. J. Appl. Logic **13**, 285–315 (2015b)

28. Magnani, L.: The eco-cognitive model of abduction. Irrelevance and implausibility exculpated. J. Appl. Logic **15**, 94–129 (2016)
29. Magnani, L.: The Abductive Structure of Scientific Creativity. An Essay on the Ecology of Cognition. Springer, Cham, Switzerland (2017)
30. Magnani, L.: AlphaGo, locked strategies, and eco-cognitive openness. Philosophies **4**(1), 8 (2019a)
31. Magnani, L.: Errors of reasoning exculpated: Naturalizing the logic of abduction. In: Gabbay, D., Magnani, L., Park, W., Pietarinen, A. (eds.) Natural Arguments. A Tribute to John Woods, pp. 269–308. College Publications, London (2019b)
32. Manders, K.: The Euclidean diagram. In: Mancosu, P. (ed.) Philosophy of Mathematical Practice, pp. 112–183. Clarendon Press, Oxford/New York (2008)
33. Marion, M.: Why play logical games? In: Majer, O., Marion, M., Pietarinen, A., Tulenheimo, T. (eds.) Games: Unifying Logic, Language, and Philosophy, pp. 3–26. Springer, Heidelberg/Berlin (2009)
34. Mumma, J.: Proofs, pictures, and Euclid. Synthese **175**, 255–287 (2010)
35. Odifreddi, P.: Divertimento geometrico. Bollati Boringhieri, Turin (2003)
36. Otte, M.: Proof-analysis and continuity. Found. Sci. **11**, 121–155 (2006)
37. Park, W.: On classifying abduction. J. Appl. Logic **13**, 215–238 (2015)
38. Peirce, C.S.: Pragmatism as a Principle and Method of Right Thinking. In: Turrisi, P.A. (ed.), The 1903 Harvard Lectures on Pragmatism. [Peirce,C.S.: Lectures on Pragmatism,Cambridge, MA, 26Mar–17May 1903, reprinted in (Peirce, 1992–1998, II, pp. 133–241)]. State University of New York Press, Albany, NY (1997)
39. Peirce, C.S.: Collected Papers of Charles Sanders Peirce, vols. 1–6, Hartshorne, C., Weiss, P. (eds.) vols. 7–8; Burks, A.W. (ed.) (1931–1958). Harvard University Press, Cambridge, MA (1866–1913)
40. Peirce, C.S.: The Essential Peirce. Selected Philosophical Writings. Vol. 1 (1867–1893). In: Houser, N., Kloesel, C. (eds.), Vol. 2 (1893–1913) The Peirce Edition Project. Indiana University Press, Bloomington (1992–1998)
41. Phillips, J.: Aristotle's abduction: the institution of frontiers. Oxf. Literary Rev. 14(1–2):171–196 (1992). Special Issue on "Frontiers", G. Bennington, G., Stocker, B. (1992)
42. Pietarinen, A.: Abduction and diagrams. Logic J. IGPL. Special Issue on "Formal Representations in Model-Based Reasoning and Abduction", Nepomuceno, A., Magnani, L., Salguero, F. Barés-Gómez, C. Fontaine, M. (eds.) (2020). Online first: https://doi.org/10.1093/jigpal/jzz034
43. Plato: Plato in Twelve Volumes. vol. I, Apology. Crito. Phaedo. Phaedrus, with an English translation by Harold North Fowler; Introduction by W.R.M. Lamb. Cambridge, MA; London: Harvard University Press; William Heinemann Ltd (1966)
44. Plato: Plato in Twelve Volumes. vol. III, Laches, Protagoras, Meno, Euthydemus, with an English translation by W. R. M. Lamb. Cambridge, MA; London: Harvard University Press; William Heinemann Ltd (1967)
45. Playfair, J.: Works of John Playfair. A. Constable & Co., Edinburgh (1882)
46. Proclus Diadochus: In PrimumEuclidis ElementorumlibrumCommentarii. ex recognitione G. Friedlein, translated and edited by G. R.Morrow A Commentary on the First Book of Euclid's Elements, Princenton University Press, Princeton, 1970. Leipzig: B. G. Teubner (1873)
47. Saito, K., Sidoli, N.: The function of diorism in ancient Greek analysis. Historia Mathematica **37**, 579–614 (2010)
48. Simson, R.: A Treatise Concerning Porisms. Simmons and Kirby, Canterbury (1777)
49. Woods, J.: Against fictionalism. In: Magnani, L. (ed.) Model-Based Reasoning in Science and Technology. Theoretical and Cognitive Issues, pp. 9–42. Springer, Heidelberg/Berlin (2013a)
50. Woods, J.: Ancestor worship in the logic of games: How foundational were Aristotle's contributions? In: Marion, M., Pietarinen, A. (eds.) The Baltic Year Book of Cognition, Logic and Communication, pp. 1–38 Volume 8, Games, Game Theory and Game Semantics, Online. New Prairie Press, Riga (2013b)

51. Woods, J.: Errors of Reasoning. Naturalizing the Logic of Inference. College Publications, London (2013c)
52. Woods, J.: Aristotle’s Earlier Logic. Second revised edition. Originally published by Hermes Science Publications, Oxford, 2001. College Publications, London (2014)
53. Woods, J.: Reorienting the Logic of Abduction. In: Magnani, L., Bertolotti, T. (eds.) Springer Handbook of Model-Based Science, pp. 137–150. Springer, Cham (2017)
54. Woods, J.: Truth in Fiction: Rethinking Its Logic. Springer, Cham (2018)
55. Woods, J., Hansen, H.V.: Hintikka on Aristotle’s Fallacies. Synthese **113**(2), 217–239 (1997)

Methodeutic of Abduction

Francesco Bellucci and Ahti-Veikko Pietarinen

Abstract Peirce's claims that methodeutic "concerns abduction alone" and that "pragmatism contributes to the security of reasoning but hardly to its uberty" are explained. They match as soon as a third claim is taken into account, namely that "pragmatism is the logic of abduction," not of deduction or induction. Since methodeutic concerns abduction and not deduction or induction, it follows that pragmatism is a maxim of methodeutic. Then, since pragmatism contributes to the security of reasoning but not to its uberty, it follows that methodeutic contributes to the security of the only reasoning it is concerned with, namely abduction. We then explain two related issues of methodeutic of abduction. First, in addition to the maxim of pragmatism, which suggests how to choose among experimental hypotheses contributing to the security of reasoning, there is the maxim of simplicity, which suggests hypotheses that are preferable for investment and which contributes to *uberty* of reasoning. Second, a third maxim of abduction is economy, which suggests adopting hypotheses that contribute to the advantageousness of reasoning even when pragmatism and simplicity cease to apply. These three maxims—experientiality for security, simplicity for uberty, and economy for advantageousness—are the bedrocks of Peirce's methodeutic of abduction.

F. Bellucci (✉)
University of Bologna, Bologna, Italy
e-mail: bellucci.francesco@gmail.com

A.-V. Pietarinen
Tallinn University of Technology, Tallinn, Estonia
e-mail: ahti.pietarinen@gmail.com

Research University Higher School of Economics, Moscow, Russia

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_5

1 Introduction

In 1902 Peirce applied to the Carnegie Institution to bring his *Minute Logic*, of which he had just written the first four chapters, to completion.[1] The so-called "Carnegie Application"[2] contained a list and the abstracts of thirty-six "memoirs," namely chapters into which his projected logic book would be divided. The last seven, Memoirs 27 to 33, were devoted to "Methodeutic." The application was infamously rejected,[3] and despite encompassing over 2.500 manuscript and typescript sheets, the *Minute Logic* itself was destined to remain severely incomplete.

Methodeutic is the name given by Peirce to the third department of his logic, after the medieval *trivium* of *artes sermocinales*: *grammatica*, *dialectica*, and *rhetorica*. The first department he called "speculative grammar," and was intended as an analysis and classification of signs. The second department, "logical critic" or simply "critic," was intended as an analysis of the different kinds of arguments (deductive, inductive, abductive). The third department, initially named "speculative rhetoric" and then "methodeutic," was described in Memoir 27 of the "Carnegie Application" as follows:

> The first business of this memoir is to show the precise nature of methodeutic; how it differs from critic; how, although it considers, not what is admissible, but what is advantageous, it is nevertheless a purely theoretical study, and not an art; how it is, from the most strictly theoretical point of view, an absolutely essential and distinct department of logical inquiry; and how, upon the other hand, it is readily made useful to a researcher into any science, even mathematics itself. It strongly resembles the purely mathematical part of political economy, which is also a theoretical study of advantages. Of the different classes of arguments, abductions are the only ones in which, after they have been admitted to be just, it still remains to inquire whether they are advantageous. (R L 75, pp. 378–379)

The idea that of the three kinds of arguments abduction is the only one in need of a methodeutic is expressed even more clearly in some drafts of the same passage:

> Methodeutic has a special interest in abduction, or the inference which starts a scientific hypothesis. For it is not sufficient that a hypothesis should be a justifiable one. Any hypothesis

[1] The preserved chapters and their sections are as follows. Chapter 1 (R 425, 176 ms pages): Intended Characters of this Treatise. § 1. Logic's Promises. § 2. Of Minute Accuracy. § 3. Different Methods in Logic. § 4. Synopsis of Contents of this Book; Chapter 2: Prelogical Notions. § 1. Classification of the Sciences (R 426–427, 291 ms pages), § 2. Why Study Logic? (R 428, 128 ms pages); Chapter 3. § 1. The Essence of Mathematics (R 431, 200 ms pages), § 2. Division of Pure Mathematics (R 430, 108 ms pages)/Specimens of Mathematical Reasoning (R 430, 68 ms pages); Chapter 4: Ethics (R 432–434, 234 ms pages). According to the plan and its revisions ("List of Proposed Memoirs on Minute Logic," R 1578), at least 33 chapters were to be accomplished in total, together with a lucrative pay-check from Peirce's friend Francis Lathrop, who had promised to pay $150 for each chapter. According to this plan, the memoirs on methodeutics were by title "27: Methodeutic. Discussion of the Method which it should possess"; "28: Economics of Research"; "29: Clearness of Ideas. Discussion of Pragmatism"; "30: Methodeutic of logical division and natural classification"; "31: On the Conduct of Inquiries and the Economics of Research"; "32: Criticism of Methods actually in use"; and "33: Objective Logic and Metaphysics."

[2] Final version and drafts in RL 75; see [1] for an introduction to the reconstruction of this set of documents, which amounts to over 500 manuscript pages in the Harvard Peirce Papers (see also NEM IV, pp. 13–73).

[3] See [2], pp. 278–289.

> which explains the facts is justified critically. But among justifiable hypotheses we have to select that one which is suitable for being tested by experiment. There is no such need of a subsequent choice after drawing deductive and inductive conclusions. (RL 75, pp. 279–280)
>
> I here consider precisely what methodeutic is. I show that it is here permissible to resort to certain methods not admissible in stechiologic [i.e., speculative grammar] or in critic. Primarily, methodeutic is nothing but heuretic and concerns abduction alone. (RL 75, p. 329)

On Peirce's justification of abduction, see [3]. In this paper, we focus on the methodeutic of abduction. We elucidate Peirce's puzzling claim that methodeutic "concerns abduction alone." As it will become apparent in the sequel, this claim is closely connected with another claim of Peirce's that needs an explanation, namely the claim, repeatedly made in 1913, that "Pragmatism contributes to the security of reasoning but hardly to its uberty" (R 838, p. 10). The two claims can be bridged, we shall argue, as soon as another claim, this time one exhaustively explicated by Peirce, is taken into account: the claim that "pragmatism is the logic of abduction." Putting these three claims together we have that since pragmatism is the logic of abduction *and not of deduction or induction*, and since methodeutic concerns abduction *and not deduction or induction*, it will follow that pragmatism is a maxim of methodeutic. Then, since pragmatism contributes to the security of reasoning *but not to its uberty*, it will follow that methodeutic contributes to the security of the only reasoning it is concerned with, namely abduction.

In the next sections, we also explain two related issues of the methodeutic of abduction. First, in addition to the maxim of pragmatism, which suggests how to choose among experimental hypotheses, there is the maxim of simplicity, which suggests hypotheses that are preferable for further investment and which contribute to the *uberty* of reasoning. Second, there is a third maxim of abduction, namely the maxim of economy, which suggests adopting hypotheses that contribute to the *advantageousness* of reasoning even when pragmatism and simplicity would cease to apply. These three maxims (experientiality for security, simplicity for uberty, and economy for advantageousness) are the three bedrocks of Peirce's methodeutic of abduction.

2 The Search for Methodeutic

It is quite late in his life that Peirce truly embarks on the study of the third compartment of logic. At some point he even seems to have taken "speculative rhetoric" (i.e., methodeutic) to comprise a study of "applications of logic to mathematics" (R 339, LN [145r], September 29, 1898).[4] What this should mean is far from clear. We know what it meant for Peirce to apply mathematics to logic: it meant using mathematical techniques and methods, for example the algebraical method, in order to solve

[4]All mathematical reasoning, he states a few years later, "relates to some schema of the nature of a diagram, that is, a sign having parts related similarly to the objects denoted, and having letters or other indices to distinguish those parts" (R 87, c.1905). Soon after, he needs to revisit these and other earlier statements of his, rethink, and widen the scope of methodeutic.

problems in logic. His work in deductive logic is an investigation into mathematical logic, and, properly speaking, is part of mathematics, not of logic. We also know that at some point Peirce thought he could give a logical account of the axioms for the natural numbers, like in the work of Richard Dedekind and Giuseppe Peano.[5] But this is hardly a matter of methodeutic, and does not explain in what sense methodeutic should be (or contain) a study of application of logic to mathematics.

In October 1906 Peirce reports to William James how during 1906 he had "attained a clearer understanding of my method of classifying relations, signs, and arguments," so much so that he took his National Academy of Science presentation approaching within only a few weeks, to be "comparatively plain sailing" (CSP to WJ, 31 October 1906; RL 224). The "method of classifying relations" Peirce mentions here may be identified with the phaneroscopical method, which Peirce had introduced in his classification of the sciences in 1902 and which he thought would constitute the basis for logic.[6] The "method of classifying signs" is, of course, speculative grammar, and the period 1903–1906 is in fact the most intense in Peirce's later production as far as the classification of signs is concerned.[7] Finally, "the method of classifying arguments" is logical critic, the sort of investigation he initiated in 1865 with the Aristotelian syllogism, and which still in 1914 would form the principal subject of Peirce's last known manuscript on logic.[8]

What he perceived still in late 1906 to be done was to complete the theory of inquiry and scientific thought, which he had famously written much about since the 1860s, with a concluding chapter on the principles of methodeutic. His studies in methodeutic, however, even more than his grammatical and critical investigations, remained largely unpublished, fragmentary and scattered across the manuscripts. On December 25, 1909, Peirce summarises his efforts and achievements to James as follows: "I have done a lot of work in Methodeutic that is valuable and very little of it is printed. This will be the most widely useful part of my Big Book. A few scraps in my 'Amazing Mazes' have brought me several admiring letters from mathematicians of some power" (RL 224). Perhaps it was when back in Arisbe, Milford, in July 1907, that Peirce really began writing the "Amazing Mazes" series for *The Monist* [8, 9] of which at least 750 draft sheets and copy-texts were completed as currently preserved in the Houghton Peirce Papers.

The "Amazing Mazes" seek to explain the role of non-mechanical, or as Peirce's choice of the word goes, "*theoric inferences*" involved in mathematical reasoning

---

[5]See [4]. In R 70(s) (9 August 1906) Peirce defines properties of positive integers in the language of Beta graphs by (i) the two-place successor function (namely that each positive integer N has a successor S(N)), (ii) that successor to be unique, and that zero is not a successor of any positive integer, and (iii) the principle of finite induction. He then derives further axioms such as that being a successor of two positive integers implies the identity of those integers, and the property that 0 is the only integer (natural number) not a successor of any integer, among other properties.

[6]On Peirce's classification of the sciences in 1902–1903 see the beautiful [5]; cf. also [6].

[7]See [7], Chaps. 7 and 8.

[8]Namely R 752 (15 March 1914). In it, he talks about "orders" of reasoning, which signals their double nature of being both "stages" and "kinds" (compare these with the terms "grades" in R 843, 1908, and "types" in R 905, 1908).

and proofs. (The phrase is misspelled in NEM III(1), p. 622 as "*theoretic* inference"; in the papers published in *The Monist* the choice of term is "theoric *step.*") Peirce wants to assure the reader that what is meant by "theoric" is a real mode of inference and not some singular and isolated creative step, element or moment encountered in the course of proofs. Peirce observes that as the members of the cyclic system grow in number, so does the size of the class of relations over those members. But the growth is enormous, factorial growth, which suggests that problems concerning cyclic systems are in the class of NP-complete problems and have the complexity equal to that of the Travelling Salesman Problem. Logical reasoning is quasi-mechanical, as mathematical proofs involve theoric inferences and as problems like those involving finding Hamiltonian cycles are fundamentally intractable.[9]

These are all pertinent observations, but they do not yet give a chapter on the scientific principles of the theory of methodeutic. Some months earlier, in March 1907, while having lodged at Prescott Hall since December 1906, Peirce was working on another long article and letter to *The Nation* and *Atlantic Monthly* on pragmatism (R 318)—and in his own words, "at a killing rate" (CSP to WJ, March 1907; RL 224). At that time, he was also preparing for the Harvard Philosophy Club lectures on "Logical Methodeutic," to be delivered in early April. Indeed these episodes almost took his life (see [2] and the appendix below for some testimonies). For what we can tell from his lecture notes, Peirce's sustained plight that spring made the delivery of these two or three talks a particularly onerous undertaking. They were to remain his last lectures, and they were presented at the Emerson Hall before the Harvard Philosophy Club on April 8 and 12, 1907, as reported by *The Harvard Crimson*: "Philosophical Club. Logical Methodeutic. I. Retroduction, or the Framing of Hypotheses. Mr. C. S. Peirce. Emerson B, 8 pm" (*The Harvard Crimson*, April 6, 1907); "Philosophical Club. Logical Methodeutic. II. Induction, or the Experimental Method. Mr. C. S. Peirce. Emerson B, 8 pm" (*The Harvard Crimson*, April 12, 1907). [10] mentions that Peirce delivered three lectures the second of which was on deduction, but there is no record of that lecture having been announced in *The Harvard Crimson*. Its draft appears to be R 754, "Second Talk to the Philosophy Club. Deduction." The draft of the lecture on induction designates it as the "Third lecture on methodeutic" (R 773), which Peirce struggled to deliver, if at all, having "been really very ill for 24 h previously and had had no sleep" (*ibid.*).

There is no surviving document for the first lecture on retroduction (abduction), but the two other fragments on deduction and induction do refer to it (Appendix).[10]

---

[9] Reasonably assuming that $P \neq NP$.

[10] R 753 appears to be an earlier, c.1906, segment of a draft lecture on retroduction, with manuscript pages 3–8 preserved. Its paragraphs that concern retroduction read as follows: "Now without taking the audience more thoroughly into this matter, I, who have gone through the whole matter far more systematically, can point out some characteristic features of these three modes of reasoning which will show their entire unlikeness to one another and what each depends upon. ¶ First, *Retroductive* reasoning is the only one of the three which produces any new idea. It originates a theory. Now if the multitude of possible theories to account for any collection of facts be not strictly infinite, it may be very moderately estimated as a million. Of these only one can be the true one. Therefore, considering that the testing of one theory is a matter that may cost a million dollars, several lives,

Perhaps Peirce was forced to compress his material on the logical methodeutic of the three types of reasoning into two lectures. Nothing is known about the reception of those talks or their attendance, save for a question from Edward V. Huntington, professor of mathematics at Harvard, which concerned the nature of mathematical discovery the answer of which Peirce reflects upon in his post-lecture log (Appendix).

3 Uberty and Security

In the last years of his life, Peirce returned to the problem of logical critic, namely to the problem of determining the degree of strength and the kind of assurance of each of the three great classes or orders of reasoning. This problem is addressed in a number of manuscripts (R 680–684, 1913), some of which discuss two distinct "poles" of the analysis of reasoning, which now Peirce calls its "security" and "uberty." Peirce goes as far as to suggest defining "logic" as a "technical term" "for the science of the security and uberty of reasonings" (R 683).

Here is an explanation of these two terms that we find amidst a long letter of 1913 to Frederik Adam Woods:

> I think logicians should have two principle aims: First, to bring out the amount and kind of *security* (approach to certainty) of each kind of reasoning, and second, to bring out the possible and esperable uberty, or value in productiveness, of each kind.
>
> I have always, since early in the sixties, recognized three different types of reasoning, viz.: First, *Deduction* which depends on our confidence in our ability to analyze the meanings of the signs in or by which we think; second, *Induction*, which depends upon our confidence that a run of one kind of experience will not be changed or cease without some indication before it ceases; and third, *Retroduction*, or Hypothetic Inference, which depends on our hope, sooner or later, to guess at the conditions under which a given kind of phenomena will present itself.

---

and a full generation, it is easy to see that history is not long enough to account for our having reached any tolerably approximate theories, without we suppose that the human mind has a natural talent for guessing at the truth of nature – meaning by nature, not only what is artificial, but all that is not downright fictitious[suffix or word(s) torn out] a product of the theory. Of course, I do not mean [word(s) possibly torn out] that a guess is probably right, but merely that it [word(s) possibly torn out] be true, to be *evident*. Being evident, it is not doubted. Not being doubted, it cannot really be criticized. For a doubt is a real state of dissatisfaction; and the common practice of making believe to doubt, and then offering considerations to appease that make-believe doubt, is a foolish waste of time, since the man who does not doubt can realize the state of mind of the man who does, only to a very limited extent, and he thus quite fails in all difficult cases to appease any real doubt that may exist, or appeases it quite otherwise than by his attempted reasonings. Another man's doubt will usually set up a doubt in my mind and that I can handle intelligently; but when his doubt excites no doubt in my mind, it is either that I know by my own experience where his mind has probably stumbled (and if I do he can be set right in a moment) or it is that he has thought that I have never had, and probably understands the matter better than I. But what appears to be *evident* is past all possible genuine criticism, until it ceases to appear evident. ¶ It will be observed that this justification of retroduction supposes that the hypothesis or conjecture which is justified has been the result of some normal way of forming such a conjecture, though the justification does not depend upon what that process is" (R 753, pp. 3–5).

> Each of these three types occurs in different forms requiring special studies.
>
> From the first type to the third the security decreases greatly, while the uberty as greatly increases. […]
>
> I don't think the adoption of a hypothesis on probation can properly be called induction; and yet it is *reasoning* and though its *security* is low, its *uberty* is high. (CSP to FAW, 6–11 November, 1913, RL 477; CP 8.384)

Deduction is the most secure of the three kinds of reasoning, for it is certain (though not merely necessary) as a kind of reasoning can be. Yet, it is the least uberous, that is, the least productive of new knowledge, because deduction concludes to nothing that was not already virtually included in the premises. Abduction, by contrast, is the least secure, because its conclusion only contains a hypothesis that has to be tested; but it is the most uberous, because it is the only kind of reasoning that can initiate a new idea.[11] As Peirce had written to the Italian philosopher Mario Calderoni a few years earlier, at a time when the terminology of "uberty" and "security" had not yet been introduced, through abduction "we gain new ideas; but there is no force in the reasoning" (CP 8.209): it "furnishes all our ideas concerning real things, beyond what are given in perception, but is mere conjecture, without probative force" (*ibid.*). Abduction is uberous because it furnishes every information about real things that is not gained in observation, but is unsecure because it has not the probative force of deduction. Naturally, in scientific investigation uberty has to be favored over security. So he wrote to Josiah Royce in 1913: "I am going to insist upon the superiority of Uberty over Security in the sense in which *gold* is more useful than *iron*, though the latter is more useful in some respects. And also that the art of making explanatory hypotheses is the supreme branch of logic" (CSP to Royce, 30 June 1913).[12]

High security and high uberty cannot both be had at the same time: "the more gravid in life and richer in light the conclusion of a reasoning may be, if it is true, the less secure can that reasoning be" (R 684, p. 2, 26 August 1913). There is a trade-off between minimizing one's errors and blunders of reasoning and increasing the value

[11] Of the three types of reasoning, abduction is "persuasive, seductive" (R 754, 1907). We are "compelled to begin with [it] if we are ever to discover a law or the rationale of any phenomenon" (R 843, p. 41, "Neglected Argument," (copy-text?) draft, crossed out, c.1907). Peirce talks about "irresistible persuasiveness" that abductive conclusions have in scientific discovery: "You will find that our rational instinct often prompts us to reasoning such that no conceivable mass of similar data would render its conclusion either certain or, in the strict sense above defined, probable, and it appears to be evident on examination that it is impossible absolutely to prove that these arguments have any value whatever. Nevertheless, it seems that many of them have an almost, if not quite, irresistible persuasiveness, that many of them have caused great discoveries and apparent great advances in science; and finally the most decisive circumstance of all in their favour is that unless these arguments have do tend [*sic*] to carry us toward new truth in the whole, we must abandon all hope of penetrating further into the secrets of the universe than we have done already" (R 652, pp. 12–13, July 12, 1910). Peirce draws an analogy from the game of Whist, in which players may be led to situations which "full warrant a player for acting on the hopeful hypothesis" (*ibid.*, p. 14). This is abduction, namely "reasoning from consequent to antecedence"; reasoning "which from a consequent and a consequence infers an antecedent" (*ibid.*, p. 15). It is "infers" that Peirce accentuates: abduction is inverse *inference* and has a certain logical form.

[12] Letter preserved in the Harvard University Archives.

of novel information.[13] Although abduction is an uberous mode of reasoning, it is "greatly wanting in security," Peirce explains in one of the 1913 papers on "Uberty and Security" (R 684, p. 3). In the same manuscript he adds that the term "'uberty' covers two quite distinct virtues" (*ibid.*), but what follows in the text unfortunately trails off and does not proceed explaining what the "two quite distinct virtues" were that Peirce had in mind.

In another draft of the same paper (R 682, reproduced in EP 2), Peirce also specifies that "uberty" is not the same as "fruitfulness":

> I can hardly be supposed to have selected the unusual word "uberty" instead of "fruitfulness" merely because it is spelled with half as many letters. Observations may be as *fruitful* as you will, but they cannot be said to be *gravid* with young truth in the sense in which reasoning may be, not because of the nature of the subject it considers, but because of the manner in which it is supported by the ratiocinative instinct. (R 682, EP 2, p. 472)

Science is the combination of observation and reasoning. It is constituted by the "processes of collecting and grouping results of Observation and of Reasoning [...] so that Science itself, when this word is used in the sense of that sort of *information* that it is the function of men of science to supply to practical men, will consist in what those men have concluded from their reasonings about observations" (R 682, EP 2, p. 471). Now, observation itself is probably the most fruitful part of the scientific enterprise, in the sense that much of the information of which a scientific theory is made comes in observation. But abduction—one of the modes of reasoning through which a scientific theory is generated on the basis of observation—is not simply fruitful: it is "uberous." Uberty, then, seems to be the kind of informativeness that is peculiar to abductive reasoning, as opposed to the informativeness that is peculiar to empirical observation, which Peirce here calls "fruitfulness." As is suggested in the passages from the 1905 letter to Calderoni quoted above, abduction is uberous in the sense that it *adds* information to the information provided by observation: it is not simply fruitful, i.e. informative in the manner of empirical observation; it is uberous, i.e. informative in the manner of reasoning.[14]

---

[13]It is only later in the process that "we correct the errors of our Retroductions [abductions] by processes of Adduction [induction]" (R 764).

[14]Here we may offer also another, and related, explanation of why uberty is not mere "fruitfulness." Peirce explains that non-uberous hypotheses can be taken to be as fruitful as we like and yet be conclusively overturned as soon as compelling evidence emerges that indicates refutation. True and false hypotheses may be both fruitful but only the former can be truly uberous. In contemporary terms, Peirce may be seen to emphasise the importance of *non-epistemic* values in science, which are not in fact distinct from *epistemic* values [11]. The uberty of retroductive conclusions suggests that non-epistemic values can be logically analysed, as he notices in the letter to F. A. Woods quoted above: "I think logicians should have two principle aims: First, to bring out the amount and kind of *security* (approach to certainty) of each kind of reasoning, and second, to bring out the possible and esperable *uberty*, or value in productiveness, of each kind" (RL 477). It needs to be conceivable that those values can be logically analysed, at least in principle. One way may be to apply dynamic modal logics of conjecture-making, which is a theory of abductive reasoning at the level of pre-beliefs, [12]. There is thus a connection between uberty and human hope that things will turn out the way anticipated: it springs from uberty that the searching questions of science upon which our rational hopes are built are amenable to final decisions. These two explanations are related, since uberty of

It is in the context of these—scanty, to be honest—remarks on the security-uberty divide that Peirce makes the following claim about pragmatism. He writes to Royce in the same late letter quoted above:

> As for my *Pragmatism*, then, it is all very well as far as it goes, it chiefly goes to improve the *security* of inference without touching, what is for more important, its *Uberty*. It does not for instance seem to have anything to say as to our exaltation of *beauty*, *duty*, or *truth*. (CSP to Royce, 30 June 1913).

The same is affirmed in another manuscript from 1913: "Pragmatism contributes to the security of reasoning but hardly to its uberty" (R 838, p. 10). In order to understand these claims, it is necessary, as anticipated in the introduction, to go back to Peirce's 1903 Harvard Lectures, where pragmatism is explicitly characterized as the "logic of abduction."

4 Pragmatism as the Logic of Abduction

In the last of his seven Harvard Lectures of 1903, delivered on 14 May, Peirce famously declared:

> If you carefully consider the question of pragmatism you will see that it is nothing else than the question of the logic of abduction. That is, pragmatism proposes a certain maxim which, if sound, must render needless any further rule as to the admissibility of hypotheses to rank as hypotheses, that is to say, as explanations of phenomena held as hopeful suggestions; and furthermore, this is *all* that the maxim of pragmatism really pretends to do, at least so far as it is confined to logic, and is not understood as a proposition in psychology. (EP 2, p. 234)

The maxim of pragmatism, in its original enunciation, states that the conception of an object is the conception of the effects of that object that might conceivably have practical, i.e. experienceable, consequences (W3, p. 266, 1878). Peirce expresses this in "What Pragmatism Is" of 1905 as follows: "[I]f one can define accurately all the conceivable experimental phenomena which the affirmation or denial of a concept could imply, one will have therein a complete definition of the concept, and *there is absolutely nothing more in it*" (CP 5.412). But the maxim, as Peirce now in 1903 sees it, is in fact a rule or test for the admissibility of hypotheses to rank as hypotheses. As the whole content of a concept consists in its experienceable consequences, so does the whole content of a hypothesis.

Abduction is the first step of inquiry, by means of which a hypothesis is suggested which, if true, would explain some surprising phenomena. The second step of inquiry is to trace necessary, or deductive, consequences from the hypothesis. These consequences are experimental predictions from the hypothesis, selected independently of whether or not they are known to be true. The third stage of inquiry, then, consists in the testing of the hypothesis through a testing of those predictions, and this is

---

reasoning has a link with truth: all good reasoning rests on logic. The defence of the latter is not to be taken up here.

induction. It consists in considering the predictions from the hypothesis, remarking what conditions should be satisfied in order for those predictions to be fulfilled, causing those conditions to be satisfied by experiment, and noting the result of the experiment. If the predictions are fulfilled, the hypothesis is inductively conferred a certain value of confidence or credence.

Pragmatism gives a criterion for the admissibility of the hypotheses that are to be subjected to inductive testing. On this, Peirce writes to Calderoni: "Pragmaticism is simply the doctrine that the inductive method is the only essential [*sic*] to the ascertainment of the intellectual purport of any symbol" (CP 8.209). Since the maxim of pragmatism states that only an experimental hypothesis can rank as a hypothesis, and since a hypothesis is experimental in as much as its deductive consequences can be verified by inductive experimentation, then the maxim of pragmatism in fact amounts to stating that the content of a hypothesis coincides with its inductively experimental content, namely with what can be inductively tested of it in experimentation.

If, as Peirce claims, *all* that the maxim of pragmatism pretends to do is to exclude non-experimental hypotheses suggested by abduction, then it has no real influence on the two other kinds of reasoning, deduction and induction, that is not already contained in its influence on abduction. Peirce explains:

> [T]he maxim of pragmatism, if true, fully *covers* the entire logic of abduction. It remains to inquire whether this maxim may not have some *further* logical effect. If so, it must in some way affect inductive or deductive inference. But that pragmatism cannot interfere with induction is evident; because induction simply teaches us what we have to expect as a result of experimentation, and it is plain that any such expectation *may* conceivably concern practical conduct. In a certain sense it *must* affect *deduction.* Anything which gives a rule to abduction and so puts a limit upon admissible hypotheses will cut down the *premisses* of deduction, and thereby will render a *reductio ad absurdum* and other equivalent forms of deduction possible which would not otherwise have been possible. But here three remarks may be made. First, to affect the *premisses* of deduction is not to affect the logic of deduction. [...] Secondly, no effect of pragmatism which *is consequent upon its effect on abduction* can go to show that pragmatism is anything more than a doctrine concerning the logic of abduction. Thirdly, if pragmatism is the doctrine that every conception is a conception of conceivable practical effects, it makes conception reach far beyond the practical. It allows any flight of imagination, provided this imagination ultimately alights upon a possible practical effect, and thus many hypotheses may seem at first glance to be excluded by the pragmatical maxim that are not really so excluded. (EP 2, p. 235)

The maxim of pragmatism excludes non-experimental hypotheses from the course of scientific inquiry. A hypothesis is experimental if its deductive consequences can be experimentally tested by induction. Thus, the influence that the maxim has on induction is already contained in its being a test for what hypotheses count as experimental hypotheses. In other words, the maxim already filters the hypotheses that will be "passed" to inductive testing, and thus it has already exhausted its task before those hypotheses are tested.

A parallel consideration holds for deduction, on which Peirce offers three points. The hypothesis suggested by abduction forms the premise of the subsequent deductive step. If the hypothesis is true (premise), then such-and-such consequences will be observable (conclusion). Thus, in the first place, the maxim of pragmatism influences deduction in the sense that it influences the premises of deduction but does not influence its "logic," that is, its inferential rationale. The second is more general, and in some sense holds for both deduction and induction: the whole influence of the maxim of pragmatism on the entire course of scientific inquiry is exhausted in its influence on abduction, so that no effect of the maxim which depends on its effect on abduction can show that the maxim is effective beyond abduction. The third point serves to block a possible objection: since deduction, on Peircean principles, consists in the *imaginative* study of the necessary consequences of wholly hypothetical states of things, how can the maxim of pragmatism influence deduction, even if only indirectly by means of its influence on abduction? Peirce's reply to this objection is that the maxim does not exclude imagination as such, but only imagination that does not "ultimately alight upon a possible practical effect" (see also [13]). If a hypothesis has in some sense a consequence that can be an object of experience, even in some recondite and indirect manner, then such hypothesis does pass the test of the maxim of pragmatism.

Pragmatism is thus exclusively the logic of abduction, and has effects on deduction and induction only through its effects on abduction. Now, given that this is Peirce's position in the Harvard Lectures of 1903, how do we connect it with his later claim that pragmatism "contributes to the security of reasoning but hardly to its uberty" (R 838, p. 10)? In the first place, as a method for excluding non-experimental hypotheses from the course of scientific inquiry, the maxim of pragmatism can hardly be said to contribute to the "uberty" of scientific reasoning as a whole. For uberty, we have seen, is the measure of the *informativeness* of reasoning; it is a value in the productivity (or fecundity) of an inferential mechanism. A non-experimental hypothesis can be as informative as any hypothesis can be, but this does not render it more likely to be true, or one that in Peirce's words would be "gravid with young truth."[15]

It is commonplace to take the meaningful measure of informativeness to be elimination of uncertainty. But is it not then the maximally informative modes of reasoning that leaves no uncertainty about its conclusion, namely deductive reasoning? Here we need to distinguish two different senses of informativeness: informativeness in uberty and informativeness in security. Reasoning becomes increasingly secure as uncertainty about its conclusions is eliminated. As noted, uberty decreases as security increases, so the measure of informativeness in the sense of uberty does not have to

[15]In one of his last essays, "Essay toward Improving Our Reasoning in Security and in Uberty," Peirce takes abductive conjectures to be distinguishable from normal formulations of interrogative moods in the sense that the former are "actually gravid with living and prolific truth" (R 683, p. 8). Not truth per se but "living truth": truth that might arise to our view in the case our inquiries were pushed to their utmost limits. Conjectures that are prolific in truth are also of "value in increasing knowledge" (*ibid.*). Peirce does not base his revisions of the logical schema of abduction directly on the shortcomings of his earlier, 1903 Harvard schema. For one, he is not led to acclaim that conjectures increase our knowledge. His later revisions concern how Modus Tollens, an inverse reasoning from effects to causes, can result in conclusions in the special "investigand" mood (RL 463; [14, 15]). The increase in knowledge follows from the imperative part of the investigand, because epistemic import cannot solely derive from the interrogative mood of questions in science. Questions do not cash out the added value in terms of knowledge. Methodological advances that Peirce refers to draw not from abduction as such but from the scientific values that imperative moods have towards resolution of pertinent research questions.

do with uncertainty and its elimination. The right method to measure informativity of uberty is rather the *simplicity* of the reasoning from premises to hypotheses. This simplicity is not itself a simple matter and it could be made more precise by speaking of *iconicity* instead. Then the higher the iconicity of reasoning is (as in abduction, in which it is the highest), the lower the security tends, and vice versa.[16] The conclusion from the maxim of pragmatism nevertheless is that it cannot contribute to the uberty of scientific reasoning.

On the other hand, the maxim does contribute to its security. For security, we have seen, is the measure of the *certainty* of reasoning. Fully secure reasoning is also maximally certain. Now abduction in itself is the least secure of the three kinds of reasoning. Yet the somewhat *unsecure* hypotheses suggested by abduction are rendered more *secure*, namely are subjected to testing, by deduction and induction. Deduction and induction eliminate uncertainty. It is the procedure of the inductive testing of the hypothesis by means of attesting its deductive consequences that confers security upon the initial hypothesis.[17] Therefore, since the maxim of pragmatism prescribes what hypotheses should count as hypotheses and what are to be excluded from the course of scientific inquiry, that maxim may be said to contribute to the security of reasoning. In other words, an uberous hypothesis is secured by means of its real verifiability, and therefore the maxim of pragmatism, in prescribing verifiability, contributes to the security of that hypothesis.

5 Economy

In a fragment preserved in a supplementary folder consisting of an assortment of loose pages we read: "The recommendations of an explanatory hypothesis are, 1st, verifiability; 2nd, simplicity; 3rd, economy" (R S-64, p. 60, c. 1903; cf. R 316a(s), 1903 in [16]. Verifiability is a recommendation that an explanatory hypothesis should accrue its security. Simplicity (as e.g. iconicity of reasoning) is a recommendation that the hypothesis be uberous. Economy is a third kind of recommendation and has to do with a hypothesis' virtue of being *advantageous* in a number of ways which experimentation and simplicity cannot ascertain and is hence to be adopted as a method to select among alternative explanations of surprising phenomena.

Being advantageous is studied in methodeutic. Methodeutic "shows how to conduct an inquiry. This is what a greater part of my life has been devoted to, though I base it upon Critic" (R 764). Peirce defined methodeutic to be "the study

[16]Many measures of simplicity abound, from model selection tasks by least-effort path principles and minimal description lengths to levels of symmetry and degrees of conservativity in guessing at new equations. What is common are the considerations of structural relationships in the inferential framework when moving from premises to conclusions.

[17]This assumes carrying out the entire tri-partite, repetitive cycles of reasoning: abduction, deduction and the experimental testing of the outcomes of the previous two stages of reasoning by induction. Indeed the action of induction is to conclude "from the results of [abduction and deduction] to what extent it will be safe to rely upon the hypothesis" (R 478, p. 102).

of the proper way of arranging and conducting an inquiry" (R 606, p. 17), depicting it as being "not so exact in its conclusions as is critical logic" (RL 75, 1902) and as involving "certain psychological principles" (R 633, 1909). But it is a theoretical study and not an art all the same. Methodeutic is based upon critic, and considers not what is admissible (logical *validity*) but what is advantageous (logical *economy*). Briefly put, it is a "theoretical study of advantages" (RL 75, 1902).

Abduction (retroduction) is of special interest to methodeutic, because it is the only mode of inference that can initiate a scientific hypothesis:

> *Retroductive* reasoning is the only one of the three which produces any new idea. It originates a theory. Now if the multitude of possible theories to account for any collection of facts be not strictly infinite, it may be very moderately estimated as a million. Of these only one can be the true one. Therefore, considering that the testing of one theory is a matter that may cost a million dollars, several lives, and a full generation, it is easy to see that history is not long enough to account for our having reached any tolerably approximate theories, without we suppose that the human mind has a natural talent for guessing at the truth of nature [...] Of course, I do not mean that a guess is probably right, but merely that it [word(s) torn out from the corner of the page] be true, to be *evident*. (R 753, c.1906; see above, footnote 10)

Being justifiable is not a sufficient property for hypotheses to be good for the purposes of science. Justified hypotheses may be immensely costly, hazardous, or utterly impracticable to be tested in anyone's lifetime. Something more is needed. In the Carnegie Application, Peirce wrote that, "[a]ny hypothesis which explains the facts is justified critically. But among justifiable hypotheses we have to select that one which is suitable for being tested by experiment" (RL 75, 1902). When one performs abductions of a scientific nature (which in one place Peirce terms "Scientific Retroductions"), they will "be subject to considerations of economy" (R 637). Considerations of economy are studied in methodeutic, not in logical critic.

Peirce's struggle to formulate a theory of methodeutic can be appreciated from the point of view of his remarks on the theory of economy of research, as it is the latter that appears as a more fully elaborated account than his late attempt at a theory of methodeutic. As those questions of what makes some inferences good and some bad are normative and value-laden, it is indeed helpful to address them from the point of view of the economy of research.[18]

A leading motivation that carried Peirce to the restatement of the logical schema of abduction in late 1903 was the realisation of the importance of economic factors that influence our reasoning in discovery and innovation. Those factors need to be integrated into the logic of abduction, and their reasons submitted to logical analysis. Peirce puts the case as strongly as the following passage could possibly reveal:

---

[18] Peirce's pertinent observation was that "now economy, in general, depends upon three kinds of factors; cost; the value of the thing proposed, in itself; and its effect upon other projects. Under the head of cost, if a hypothesis can be put to the test of experiment with very little expense of any kind, that should be regarded as a recommendation for giving it precedence in the inductive procedure" (CP 7.220, 1901). On Peirce's economy of research, see e.g. [17–20]. [21] has introduced related concepts of eco-cognitive openness, optimization of situatedness and eco-cognitive model of abduction and can be used in explaining further the phenomenological aspects of creativity in scientific reasoning.

> The principles upon which abduction ought to be conducted ought to be determined exclusively by considerations of what purpose it subserves and how it may best subserve that purpose. Since, therefore, in scientific investigation abduction can subserve no other purpose than economy, it follows that the rules of scientific abduction ought to be based exclusively upon the economy of research. (CP 7.220n)

Rules of abduction are based only upon what the economy of research is capable of prescribing. Now among critically equivalent hypotheses (that is, hypotheses that explain the same surprising facts), one should be able to select for testing those that are capable of experimental verification. This, as we have seen, is the core of Peirce's maxim of pragmatism,

However, there may be inconveniently many "pragmatistically" equivalent hypotheses, that is, hypotheses that are capable of experimental verification. Thus among such pragmatistically equivalent hypotheses one should select for testing those that in the sense of the economy of research are the "cheapest" ones. Peirce's argument for the economic character of methodeutic is roughly as follows: the logical validity of abduction presupposes that nature be in principle explainable. This means that to discover is simply to expedite an event that would sooner or later come to pass. Therefore, the real service of a logic of abduction is of the nature of an economy. Economy itself depends on three factors: cost (of money, time, energy, thought), the value of the hypothesis itself, and the effects that it renders upon other projects and hypotheses (RL 75, 1902; R 690, CP 7.164–231, 1901).

Of permanent importance in economic issues is the question of the allocation of resources. A closer analysis of abductive reasoning reveals the presence of it in what may be termed the Maxim of Novelty, namely that "the new money should mainly go to opening up new fields." This is, Peirce explains, because "new fields will probably be more profitable, and, at any rate, will be profitable longer" (CP 7.160). Chances of discovery are highest in frontier and interdisciplinary research, as it is in those areas that the constraints that may otherwise be imposed upon finding the right antecedents of subjunctive conditionals in the second premise of abduction are deliberately and decidedly turned down. Curiosity-driven, exploratory ideas must reign free in order to maximise chances of someone hitting upon new hypotheses that may be gravely needed. Finding novel methods and tools of analysis are equally profitable, and are according to the economy of research commended to remain part and parcel of blue sky research, because those methods are both cheap to pursue and at the same time broadly applicable across sciences, without limitations (cf. CP 7.161). In contrast, funding several competing research groups working on roughly similar problems, driven by motives such as "who gets there first," is not conducive to the real advancement of inquiry and just contributes to the increase of false discovery rates [22]. Nor is the research driven by some singular sets of values or grievances of special interest groups or identities.

Peirce distinguished three components in his argument for the economy of research. First, hypotheses possess the quality of *caution*: according to it hypotheses are to be broken down into their smallest logical components. Big questions are to be divided into series of small questions; *why*- and *how*-questions into series of *yes-no*-questions.[19] According to the quality of *breadth*, the value of a hypothesis is evaluated by its applicability in other related subjects across a multiplicity of contexts and circumstances. Explanations of the same phenomena should be evaluated according to their consequences. Finally, the quality of *incomplexity* (the absence of complexity, simplicity, or artlessness) states that, because hardly any hypothesis is optimal anyway, they ought at least to "give a good leave" (EP 2, p. 110). Refuting a null-hypothesis is not only a refutation but a setting of an example of good conduct to be followed, as investigators should attempt as large a "break" as possible from a hypothesis that is destined to fail, for contingent reasons that "we must always consider what will happen when the hypothesis proposed breaks down" (CP 7.220).

In addition, the science that studies logical methods of science is itself cheap to employ, and by virtue of the principles of the economy of research should be applied to a variety of difficult problems that otherwise would be too costly or slow to pursue. The power of signs is evaluated as they appeal to the mind. It is the question of methodeutic how to render signs *effective*. The task is to map out the many ways in which interpretants are linked to symbols, indices and icons. The task is also to find out how the illocutionary force of retroductive conclusions, conceived as assertions [23], are transmitted to the minds of their interpreters, and with what consequences. Special illocutionary forces are associated with conclusions drawn when one reasons from the consequent and consequences to new antecedents. The meaning of one's research questions is thus to be analysed in tandem with the meaning of one's hypotheses.

All these are questions that the theory of the economy of research is calculated to address. Economic considerations are thus pertinent to the development of the method of finding right kinds of questions to be asked in science. Finding good questions to ask is possibly the hardest thing in actual scientific practice. Exploring why this is so is a perpetually important topic, and some advice may be forthcoming from looking at the matter from the logical point of view; after all, according to Peirce, "it is a question whether [economics] is not a branch of logic" (RL 75).

6 Conclusion

Once critic has justified abduction, it is up to methodeutic to teach how to reason abductively in an *effective* way. Abductions are for Peirce "the only ones in which after they have been admitted to be just, it still remains to inquire whether they are advantageous" (RL 75, 1902).

In addition to the maxim of pragmatism, which suggests choosing among experimental hypotheses and which contributes to the *security* of reasoning, and to the maxim of simplicity, which suggests that those hypotheses are preferable for further

[19]The Nobel Prize in economics went in 2019 to poverty research that did not ask the unanswerable "How to get rid of poverty?" but, for instance, "Could the rate of access to clean water be expedited?", etc.

investment which are "simple" (in the Peircean–Galilean sense, see [14]) and which contributes to *uberty* of reasoning, we have a third maxim of abduction: the maxim of economy, which suggests a method for adopting hypotheses in the senses qualified above (cost, novelty and impact) and which contributes to the *advantageousness* of reasoning when pragmatism and simplicity would cease to apply. These three maxims (experientiality, simplicity and economy) are the bedrocks of Peirce's methodeutic of abduction.

Acknowledgments Supported by the Estonian Research Council personal research grant PUT 1305, Abduction in the Age of Fundamental Uncertainty (A.-V. Pietarinen, 2016-2018), and the HSE University Basic Research Program funded by the Russian Academic Excellence Project '5-100' (A.-V. Pietarinen, 2018-2020).

Appendix

R 754 [p. 1]

Second Talk to the Philosophical Club

Deduction. Define it. Several years' *reconsideration* leads to this new definition.

Not *necessary* but *compulsive* reasoning.

Retroduction persuasive, seductive.

Induction appeals to you as a reasonable being.

Deduction points to the premises and to their relation and then shakes its fist in your face and tells you: "By the eternal powers, you have *got to* admit the conclusion." *Effort required* distinguishes deduction. That [it is] not necessary [is] shown by *definition of the necessary* and argument about *gold*. [See the last paragraph of this R 754 for the argument about gold.]

Compulsive means that you are logically forced to admit the conclusion.

Now, resume the division of reasoning. You begin with *perception* which brings *surprise*. To make that reasonable you resort to Retroduction. ~~That finally finished~~ This resembles perception in bringing the *new*. And after this no new idea enters into reasoning. Now comes Deduction by which this new forces you to join with it a *transformation*,—no new *matter* but only a new *form*. And Deduction is *defined* as including all that. Finally Induction, or the Experimental Method tests its truth.

Observe *Regular Gradation* from the incoming of the new, first in Perception, then in Retroduction; then, *formally only* in Deduction, while the reasonableness, quite absent in perception begins to appear in Retroduction, [and] still more forcibly in Deduction and in its highest development in Induction.

[p. 2, discontinuous] […] the depths of consciousness. Question of number of premises. Effort required. Theôric element never or hardly ever absent from deduction.

I must confess I have nothing but such mere *Generalities* to offer, when what is wanted is a **Method for the discovery of Methods**. Utter weakness of all the Books. One obvious and old remark is that a good way to prove that a thing is possible is to show how it can be brought about. Yet that that is not indispensable is shown by a celebrated proposition in the elements of higher arithmetic.

The Mathematicians [are] the only skilled and correct reasoners because all reasoning depends on deduction and I will not consent to my admired friend Bôchler's limitation of mathematics to any particular class of Deductions. Mathematics is the practice of deduction.

It seems to me rather remarkable that though it will not be disputed that I have shown high powers in regard to the *theory* of deduction, I have but a very moderate mathematical talent. Perhaps what I mostly [blame] imperfection is my

Proof of abnumeral multitudes.

Next the proof which I was the first to give [was] that Linear Associative Algebra is the same things as the Theory of Matrices.

Obvious enough, perhaps, yet disputed by Sylvester. Then my proof that only three algebras have determinate division [quaternions, octonions, nonions].

Then various trifles [are] relative to computations in which my lack of generalizing power is most prominent. (Duplication of the cube. Entire circulating decimals.) [p. 3]

The more I consider the matter, the more I am convinced that corollarial reasonings are the highest.

Among these is my proof about multitude.

Also De Morgan's Syllogisms of Transposed Quantity.

Also the Fermatian Inference commonly but most absurdly called Mathematical Induction.

But those stupendous proofs of Gauss appear to be Theôric. That is why he can prove a proposition in so many different ways. There can be but one sound Corollarial proof of any thing from given premises, though the order of procedure may be somewhat varied. [p. 4]

Second talk. On Deduction.

It was Monday. I was overcome by this Boston April: how ill it always made me and nearly drove me frantic. Was so ill at last lecture and so gave to sundry points colors I had never intended.

I began by speaking of myself,—a very hazardous and rash thing to do;—one runs away with oneself so easily and finds oneself brought up in the court of self to answer the charge of driving beyond the limit of speed.

I mentioned my unusual turn for the subtleties of logic, but entirely forgot to mention, what I had fully intended to emphasize, the quite *adventitious* character of such gifts. A talent for logic, or any other, is no more a thing to take credit to oneself for than is the inheritance of a fortune. The parable of the talents embodies deep truths. The talents are represented as capable of growth if properly put to use. This is very true, indeed. But the parable leaves us to think that the talents are more entirely extraneous than they really are, that they are as passive as so many pieces of gold, and that the issue of the matter depends entirely upon the personal character of the man, regardless of the talents. It is true that there is an attempt, in the parable to [p.6] correct this error by representing that it is only the man who has one isolated talent who fails to put it to use, being fearful lest he should lose his all. Well, I am a Christian, as we all are, did we but recognize it. But the gospel of Christ is a gospel at once of liberty and of duty to think out things for ourselves, and we have no warrant whatever for regarding the parables as anything else than appeals to our reason. Now this talent that we receive, *even such a limited endowment as mine*, is more than a fixed sum. It is capable of development so as to dominate and elevate the whole character; and therefore I prefer that other parable or metaphor of the Nancy hypnotists of multiple personality;—an ancient and Christian [*illeg.*] it is too, of a spirit that grows within a man and supplants the "old man." [Added in margin:] It is a singular result of well-known circumstances, that a man is not permitted to express his belief in a personal God in any impassioned way, no matter how overwhelming that belief may be.

Dr. Huntington at the end of the other evening's talk asked how I would regard the great hypotheses of pure mathematics; and whether it was that I was fatigued or what, I replied quite stupidly. I certainly regard from considering what for want of a better word I may call the facts of mathematics. When a coefficient of an equation gradually changes so that two adjoining roots after coalescing [p. 5] disappear, our explanatory hypothesis to account for the phenomena if you will permit me to so to call it, is that the roots, in consequence of the collision have struck off into another dimension. It is a hypothesis, because it is a way of rendering the phenomena comprehensible. But it is only a formal hypothesis, not a material hypothesis[.] Definition or Division are also retroductive. Deal with them separately.

Deduction. I have hitherto defined this as *necessary* reasoning; and no doubt much, perhaps most, possibly all deduction *is* necessary. But on reviewing the subject for this talk, it seems to me more correct to define Deduction as compulsive reasoning. Retroduction seduces you. Induction appeals to you as a reasonable being. But Deduction first points to the premises and their relation, and then shakes its fist in your face and tells you "Now by God, you've got to admit the conclusion." I beg your pardon, with all my heart, I meant to say, "Now by the eternal world forces spiritual and personal [*illeg.*]" Necessary reasoning is reasoning from the truth of whose premises it not only follows that the conclusion *is* true, but that it would be so under all circumstances.

Gold. Somebody submits a proposal for a new coinage, with a specimen the Master of the mind says, This is pure gold and therefore it is too soft for the purpose. That is not *necessary* reasoning for it is quite conceivable that gold should be hardened as pure copper and pure iron can be. But in existing circumstances, no such way of hardening gold being known, it is now quite compulsive. [End of R 754]

Third Lecture on Methodeutic [R 774]

Induction

Apology. Had been really very ill for 24 h previously and had had no sleep. I am still far from well, but hope to do better than last night.

That proof which I failed to give is as follows. No matter what objects the Ns may be, so long as they have independent identities, I ask whether there can be any relation in which every conceivable collection of Ns stands to some N to which no other describable collection of Ns stands in the same relation.

Take any relation you please, and call it R. Then I divide the Ns into 4 classes:

First class consists of Ns to each of which no collection of Ns is R.

Second class consists of Ns to each of which 2 or more collections of Ns are R.

Third and fourth classes consist of Ns to which of which a single collection of Ns is R.

The third class consists of those of these Ns each of which is contained in their collection of Ns that is R to it. The fourth class of those each of which is uncontained in the collection of Ns that R to it.

Now I will describe a collection of [p. 2] Ns which is not R to any N, unless there be another collection of Ns that is R to it. It shall contain all the Ns of class 4 none of class 3 and [2]. This is the collection I choose as a test. Is this the sole R of any N. Not of any N of the first class, since no such has any R at all. N of any N of the second class since all such have two Rs so that [*illeg.*] is *sole* R to any of these. Nor is the collection I have described R to any N of the third class, since no N of the third class is contained in it, while every N of the third class is contained in the sole collection that is R to it. Nor is this collection R to any N of the fourth class, since every such N of [*sic.*] is contained in it, while no N of the fourth class is contained in the sole collection that is R to it. So here is a collection described as definitely as you please which is not the sole R of any N.

I spoke of Deduction as the compulsive kind of reasonings. Almost all the theoric inferences are positively *creative*. That is, they create, not existent things, but *entia rationis* which are quite as real. This blackboard is black. Theoric deduction concludes that the board possesses the quality of blackness and that *blackness* is a simple object, called [p. 3] an *ens rationis* because that theoric thought creates it.

Object and Interpretant of a Sign.

The interpretant of Deductive Reasoning is Energetic, or as I call its *generalization* Existential. The Object of it, however, is purely Hypothetic. If that situation which Deduction reasons about happens to exist, that has no bearing at all on its reasoning. It is not so with Retroduction which starts from the actually perceived. Still less is it true of Induction which as we shall see, relates entirely to the actual course of Experience as its Object.

Induction, as I use the word, and I confess I cannot defend my case of the word,— is simply the acceptance of a hypothesis in so far as it has supported tests. Every test of a hypothesis consists in making it the basis for a conditional prediction, and if the prediction turns out to be verified, we reasonably accept the hypothesis provisionally. That is the highest certainty attainable by positive science.

The justification of it is that if it be false[,] *perseverance* in the method will correct its own result. Induction has three grades. The first is the simple assumption that the future will be like the past. Hence we deny fairies, ghosts, telepathy, etc.

The second is reasoning from a random sample. [p. 4]

The third is where a random sample is impossible because no counting or measurement is possible, but where we know that such a method must bring indefinite approximation to truth in the long run.

Organic chemistry. Aristotle. Plato and his Xth letter. Basil Valentine. [End of R 774]

References

1. Ransdell, J.: The Significance of Peirce's Application to the Carnegie Institution, (1998). www.iupui.edu/~arisbe/menu/library/bycsp/L75/intro/l75intro.htm
2. Brent, J.: Peirce: A Life. 2nd edn. Bloomington & Indianapolis: Indiana University Press (1998)
3. Bellucci, F., Pietarinen, A.-V.: Peirce's justification of abduction. Stud. History Philos Sci. A. **84**, 12–19 (2020)
4. Shields, Paul, Charles, S.: Peirce on the Logic of Number, Docent Press, Boston (2012)
5. Kent, B.: Charles S. Peirce. Logic and the Classification of the Sciences. McGill-Queen's University Press, Kingston (1987)
6. Pietarinen, A.-V.: Interdisciplinarity and Peirce's classification of the sciences: a centennial reassessment. Perspect. Sci. **14**(2), 127–152 (2006)
7. Bellucci, F.: Peirce's Speculative Grammar: Logic as Semiotics. Routledge, New York (2017)
8. Peirce, C.S.: Some amazing mazes. Monist **28**(2), 227–241 (1908a). (Reprinted in CP 4.585–593)
9. Peirce, C.S.: Some amazing mazes (conclusion). Explanation of curiosity the first. Monist **28**(3), 416–464 (1908b). (Reprinted in CP 4.594–642)
10. Fisch, Max H.: Peirce's place in American life. Historica Math. **9**, 265–287 (1982)
11. Chiffi, D., Pietarinen, A.-V.: Fundamental uncertainty and values. Philosophia **45**(3), 1027–1037 (2017)
12. Ma, M., Pietarinen, A.-V.: Let us investigate! dynamic conjecture-making as the formal logic of abduction. J. Philos. Logic **47**(6), 913–945 (2018)
13. Bellucci, F., Pietarinen, A.-V.: The iconic moment: towards a Peircean theory of scientific imagination and abductive reasoning. In: Pombo, O., Nepomuceno, A., Redmond, J. (eds.) Epistemology, Knowledge, and the Impact of Interaction, pp. 463–481. Springer, Dordrecht (2016)
14. Pietarinen, A.-V., Bellucci, F.: New light on Peirce's conceptions of retroduction, deduction, and scientific reasoning. Int. Stud. Philos. Sci **28**(2), 353–373 (2014)
15. Pietarinen, A.-V.: Abduction and diagrams. Logic Journal of the IGPL. (2019 in press)
16. Peirce, C.S.: Logic of the future: writings on existential graphs. Vol. 1. History and Applications. Ed. Pietarinen, A.-V. Berlin & Boston, De Gruyter
17. Haack, S.: Expediting inquiry: Peirce's social economy of research. Trans. Charles S. Peirce Soc. **54**(2), 208–230 (2018)
18. Paavola, S.: Abduction as a logic and methodology of discovery: the importance of strategies. Found. Sci **9**, 267–283 (2004)
19. Rescher, N.: Peirce and the economy of research. Philos. Sci. **43**(1), 71–98 (1976)

20. Wible, J.R.: The economic mind of Charles Sanders Peirce. Contemp. Pragmatism **5**(2), 39–67 (2008)
21. Magnani, L.: Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning. Springer, Dordrecht (2009)
22. Ionnadis, J.: Why most published research findings are false. PLoS Med **2**(8), e124 (2005)
23. Chiffi, D., Pietarinen, A.-V.: Abductive Inference within a Pragmatic Framework. Synthese **197**, 2507–2523 (2020)
24. Chiffi, D., Proover, M., Pietarinen, A.-V.: Anticipation, abduction and the economy of research: a normative stance. Futures J Policy Plann Futures Stud **115**, 102471 (2020)
25. Peirce, C.S.: The collected papers of Charles S. Peirce. 8 vols. Eds. Hartshorne, C., Weiss, P. Burks, A.W. (eds.) Harvard University Press, Cambridge (1931–1966)
26. Peirce, C.S.: Manuscripts in the Houghton Library of Harvard University. Catalogued in Robin, R. (ed) Annotated Catalogue of the Papers of Charles S. Peirce, Amherst, University of Massachusetts Press (1967); and Robin. R. (ed) The Peirce Papers: A supplementary catalogue. Trans. C. S. Peirce Soc 7, 37–57 (1971)
27. Peirce, C.S.: The new elements of mathematics by Charles S. Peirce. 4 vols. Ed. Eisele. Mouton, The Hague (1976)
28. Peirce, C.S.: Writings of Charles S. Peirce: a chronological edition. 7 vols. Ed. Peirce Edition Project. Indiana University Press, Bloomington (1982–)
29. Peirce, C.S.: The essential Peirce. Vol. 2. Ed. Peirce Edition Project. Bloomington & Indianapolis: Indiana University Press (1998)

Peirce, Russell and Abductive Regression

John Woods

Abstract Below are reflections on Peirce's conception of abductive methods and Russell's conception of regressive methods. Along the way, it will be necessary to examine the marked differences between Russell and Frege on the ins and outs of logicism, from which latter the regressivist ideas first emerged. Russell was aware of Peirce's contributions to the algebraization of logic and Peirce was aware of Russell's writings on logicism. However, in framing his thoughts about regressive methods, Russell showed no familiarity with Peirce's treatment of abductive methods. In 1907, Russell read to the Cambridge Mathematics Club an essay entitled "The regressive method of discovering the premises of mathematics." Since that paper didn't see the published light of day until three years after Russell's death in 1970, Peirce couldn't have taken notice of it in developing his ideas about abduction. Even so, it has been suggested that there exists a noteworthy similarity between Russell's regressivism and Peirce's abductivism. The principal purpose of this essay is to show the resemblance to have been misjudged, in which case, my title would have to be corrected.

1 Peirce

1.1 Peirce's Abduction

Peirce was an algebraic logician, a principal developer of the logic of relations and the founding father of quantification theory. Credit for the latter usually goes to Frege, but Quine and Putnam are oppositely inclined.[1] So am I. It may be well to take early

[1]See Peirce [25–27]. Putnam [32, p. 297] and Quine [34, p. 259]. See oppositely (and earlier) Goldfarb [12]. We needn't settle the matter here.

J. Woods (✉)
University of British Columbia, Vancouver, BC V6T 1Z1, Canada
e-mail: john.woods@ubc.ca

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_6

notice of Peirce's disinclination to see his account of how we think when we draw abductive inferences as part of logic "in the strict sense."

> My proposition is that logic, in the strict sense of the term, has nothing to do with how you think ….[2]

> Indeed,

> it is plain that the question of whether any deductive argument, be it necessary or probable, is sound is *simply* a question of the mathematical relations between one hypothesis and another. (ibid., 144, emphasis mine.)

> Although logic is solely concerned with the presence or absence of mathematical relations between propositions, it nevertheless remains the case that propositions must be expressed somehow; and for this reason formal logic, in order to disentangle itself completely from linguistic, or psychical, considerations invents an artificial language of its own, of perfectly regular formation, and declines to consider any other form of statement then in that artificial language. (ibid., 144–145)

In matters of how we do think, and should, Peirce has a good deal to say. He was in no doubt that the ins and outs of abductive reasoning are expressible and enactable in ordinary human speech. Peirce framed his views on abduction as a contribution to the philosophy of experimental science. We could say, then, that it is an essay on the *epistemics of science.* Important as it is for human cognitive prosperity, abduction for Peirce is not a matter for logic in the strict sense.

Let's turn now to Peirce's famous schema.

1. The surprising fact F is observed.
2. But if H were true, F would be a matter of course.
3. Hence there is reason to suspect that H is true.[3]

Although Peirce doesn't specify what in line (1) he means by "surprising", "a matter of course" in line (2), or "there is reason" and "suspect" in line (3), the general idea becomes clearer when we add further details of his account. For example, abduction is a form of *guessing*, underwritten *innately* by *instinct*. (RLT, p. 128, CP5. 171, CP 7.220) Abduction provides *no grounds for believing* the abduced proposition to be true. (RLT, p. 178) Instead the correct thing to do with abduced hypotheses is send them off to *experimental trial*. (CP 5.599, CP 6.469–6.473, CP 7.202–219) The connection between an abduced hypothesis and the observed fact is *subjunctive.* (CP 5.189) The conclusion licensed by abduction is not proposition H, but rather the proposition that H's truth is a matter for reasonable *suspicion.*

[2]Peirce [28, p. 143], Hereafter RLT.

[3]Peirce (CP 5.189), in [29] Citations are structured as follows: "CP 5.189" denotes Volume V of *Collected Papers,* p. 189. Line numbers are for referential convenience, and I have changed the original schematic letters to more intuitive ones and removed the italics.

1.2 The Gabbay-Woods Approach

A *Peircian conception* of abduction is one that preserves, refines or expands upon the features presently at hand. The best-known to me of such adaptations is set out in what has come to be called the Gabbay-Woods Schema. On the Gabbay-Woods approach abduction has two separate but interlocking parts.[4] The feature that links them is a fact-created *ignorance-problem.* Often enough, one's ignorance is not problematic. Not knowing why Robbie prefers raspberries to strawberries or Maddy the other way round, is not something of any great moment to me. But there are occasions on which not knowing something is irritating, and not at all something we'd be prepared simply to give up on. So, until further notice, a better name is *ignorance-irritation* problem.

Here is the basic structure of an ignorance-irritation problem. With our present store of knowledge and the further knowledge that lies within our timely reach, we don't know how to account for this fact. Relative to these limitations, the question is *intractable.* To arrive at an answer there are two requirements to be fulfilled. One is the selection of a hypothesis H from up to arbitrarily large classes of possibilities. What is wanted is a hypothesis H which, when added to our present knowledge-base K, would be such that if the resultant K(H) held true, the questioner's ignorance would dissipate. Let's call this the *selection-step.* The problem of hypothesis-selection is made more complicated when, as frequently happens, more than one hypothesis preserves the truth of the subjunctive conditional at line (2). Sometimes, indeed, these rivals are pairwise inconsistent. As we have it so far, the hypothesis-selection stage is the least well-developed part of the Peircian approach, and is not, as we are about to see, much improved upon in the Gabbay-Woods adaptation of Peirce's schema. In addition to the rival hypotheses problem, there are also strictly logical problems when F is the terminus of a search for premisses. Suppose that some true F were the conclusion of a valid argument.[5] What, we might as, are its premisses? Does F itself count as a premiss? Does the conjunction of F and its negation qualify? Does the largest self-consistent set of truth-evaluable English sentences severally and jointly consistent with F count? These questions answer themselves and yield up a moral. There is more to being a premiss for F than being a statement or statement-set that logically implies it.

Before leaving this example, we should note that "matter of course" clause need make no appearance here, as would be evident if F were the proposition that $1 + 1 = 2$, and the question it posed were whether it is finitely provable. The same for the "surprise" clause. Ignorance-irritation problems can be triggered by completely

[4]For a recent more detailed exposition, see Woods [45]. See also Gabbay and Woods [11], Magnani [22] and Park [23]. All these works adopt naturalistic assumptions for the logic of inference. A still unsettled disagreement is about whether all abductions are inferences to the best explanation. The authors cited in this note think that it is not that *intrinsically.* But here, too, we needn't settle the question for what matters here.

[5]For expositionary ease, I use "F" ambiguously. In Peirce's Schema it names some given fact or state of affairs. When it occurs as the conclusion of a piece of reasoning, it names a proposition that states that fact. I leave it to context to disambiguate.

unsurprising facts. Equally, one's ignorance needn't be irritating. It suffices for it to be *motivating*.

The next step(s) are in fact two. One is to determine what is to be concluded from the successful completion of the selection-step. In what I'll call the *inference-step*, Peirce authorizes the conclusion that "[t]here is reason to suspect that H is true." Next, in what I'll call the *action-step*, Peirce provides that if the costs are not too high and the benefits sufficiently desirable, it is reasonable or even advisable to release H for provisional premissory use in the domain of enquiry in which the ignorance problem arose in the first place. Thus

Suspicion: The *operational meaning* of our suspicion of a hypothesis' truth is our cost-benefit willingness to submit the hypothesis to an epistemic stress test.

This sets up the fourth and final abductive stage, the *test-step*. Should the now-released H play a role in advancing our cognitive prosperity in the domain of enquiry at hand, it would achieve—usually by degrees—some positive approximation to a settled fact. The test is passed by the fruitfulness of H's employment.[6] If the test is not met, H loses its premissory license and drops from sight.

Here now is a slightly more formalized version of the Gabbay-Woods Schema. Let F be a fact that prompts a question in someone, Bill say, to which he lacks the answer he desires to have. Let F? be that question. K is Bill's current knowledge-base and K* is the immediate successor of K, made so by within Bill's timely reach. (For example, he could find out be going online.) Let H be a hypothesis and K(H) the update of Bill's knowledge-base upon addition of hypothesis H to K. Let $\looparrowright$ be the subjunctive conditional connective and R the ignorance-lifting response-relation. C(H) denotes the conjecture of H as a truth-suspect, and H^C denotes the decision to take action on H. Then the general structure of the G-W extension of Peirce's schema is:

1. F	[fact]
2. ?F	[question]
3. $\sim(R(K, ?F))$	[fact]
4. $\sim(R(K^*, ?F))$	[fact]
5. $H \notin K$	[fact]
6. $H \notin K^*$	[fact]
7. $\sim R(H, ?F)$	[fact]
8. $\sim R(K(H), ?F)$	[fact]
9. $H \looparrowright R(K(H), ?F)$	[fact]
10. H meets further conditions $S_1, \ldots, S_n$	[fact]
11. Therefore, C(H)	[subconclusion, 1–10]
12. Therefore, H^C	[conclusion, 1–11]

[6]Fruitful conjecture is nicely discussed by philosophically trained historians of science. See, for example, Tappenden [42] and Ferreirós and Gray [7].

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 Bill's knowledge-set K. Nor is it in its immediate successor H*. Since H is not in K, then the revision of K by H is not a knowledge-successor K* of K. Even so, if H were true, then it would be the case that K(H) would answer ?F. Note, however, that even though this is a successful abduction, this *itself* is not an answer to ?F. Rather it gives license to try to find out whether H in fact *is* true. For if it were true, ?F would have an ignorance-erasing answer. It is commonly accepted that the distinguishing mark of deduction is truth-preservation and, of induction, is likelihood-enhancement. If the Gabbay-Woods model is faithful to Peirce, then the distinctive feature of Peircian abduction is *ignorance-preservation,* but in a way that identifies candidates for its possible removal upon further investigation. Let us be clear. Neither of the schema's conclusions answers the question posed by E. They are helpful and instructive in other ways.

Line (10) leaves the conditions on hypothesis-selection unschematized. But it isn't true that its authors stand mute on the matter. In Woods [46], much is made of the fact that among the many cognitive tasks at which we humans are adept, is the selecting plausible or "ballpark" hypothesis in ?F-contexts. In plainer words, we are really quite good at good guessing. There are other things we are rather good at. We have just-in-time memories that generally serve us well, and we're really quite good at keeping irrelevant information at bay by denying it premissory license. We are good at spotting local inconsistencies arising from incoming information – this, after all, is the *conditio qua non* of pro-contra argument, the life's-blood of the *reductio.* On the other hand, we're no good at all in keeping *deep* memory and *background* information consistent.[7] But these frailties tend to have little impact on the premissory use of small consistent subsets of those larger wholes. A final example is the skill we display in managing the distinction between a proposition's or a theory's deductive and demonstrative closures. It is a theorem of number theory that zero is a natural number. It follows from this theorem that zero is a natural number or Nice is nice in November. "Zero is a natural number or Nice is nice in November" is in the deductive closure of Peano-Dedekind arithmetic, but it is not a theorem of arithmetic, it is not an arithmetic truth. The reason why is that it is not in the theory's *demonstrative* closure. Nobody who has ever drawn breath has identified "Zero is a natural number or Nice is nice in November' as a truth of arithmetic. That is because we're so good at managing the requisite closure differences. For the most part, we have no conscious grasp of how we manage to be good at these things. This fact alone disposes me to think that these mechanisms that filter out irrelevances and regulate vital distinctions are for the most part automatic. This is a conjecture of a type entirely welcome by the causal-response epistemologies of naturalized logic.[8] The reason that they aren't listed in line (10) of the schema is that it is not yet known with any specificity what

[7] More on this can be found in Chap. 9, "Putting inconsistency in its place", of Woods [46].

[8] A causal-*response* epistemology sees the beliefs that qualify as knowledge as causal responses to information-processing under specified conditions. (Woods [44], Sect. 3.5). In contrast, a causal-*contact* epistemology, requires a believer's causal contact with the object of belief [13].

they are and how they work. So until something better comes along, perhaps we could rewrite (10) to say:

(10′) H is the product of well-filtered possibilities.[9]

2 Russell

2.1 Mathematics When Russell Was Starting

Conjecture, of course, is the dynamic engine of all living science. Without new ideas, the sciences would atrophy. No nineteenth century mathematician would have dreamt of mounting an arrest-on-sight anti-conjecture campaign. The general question raised by conjectures is what is to be done with them when they first arise? No one would have doubted that the right thing to do with a scientific conjecture is put it to the test, put it, so to speak, on trial. In the empirical sciences, a winning conjecture will have to negotiate its way through the confirmational checkout counter. In the abstract sciences, there is no question of empirical validation. The general and universally received answer is that even an abstract conjecture wins the day if it is backed by a valid proof. Normally, however, there is a time-lag between the emergence of a conjecture and a proof that validates or invalidates it. For example, Poincaré's Conjecture arose in 1904 and wasn't proved until 2006. Riemann's Hypothesis was announced in 1859 and still lacks a settled consensus. Hodge's Conjecture was announced in 1950 and also awaits a solution.[10] It is natural to ask of conjectures in limbo—102 years in Poincaré's case, 106 years and counting in Riemann's, and 69 and counting in Hodge's—what are the rules of their employment? It is here that opinion starts to divide. Many mathematicians hold that limboed conjectures should be allowed to circulate and seek profitable careers in conditional proofs, for example; especially when they have consequences which themselves turn out to be independently provable. Opinions divide more sharply over the question of whether mathematically interesting ideas should ever be granted provisional *premissory* license to appear in categorical proofs.

At the heart of this matter are considerations of risk-tolerance. The landscape of Peirce's abductive surveyance was a nineteenth century commonplace in mathematics and science, and is so to this day. During the time of Peirce's mathematical flourishment and of Russell coming to some degree of mathematical maturity, mathematics was in a state of creative crisis. On the one hand, new ideas came "fast and furious and curious". Among the more interesting newcomers were Riemann's n-tuply extended magnitudes, the irrational, imaginary and complex numbers, elliptical functions, progressive harmonic conjugates and Desargesian planes. On the

[9]For more enriched versions see Chiffi and Pietarinen [3, 4].

[10]Poincaré's Conjecture says that every simply connected closed 3-manifold is homeographic to the 3-sphere. Riemann's Hypothesis is that all non-trivial zeros of the Riemann zeta function have a real part equal to 0.5. Hodge's Conjecture is that certain de Rham cohomology classes are algebraic.

other hand, in the midst of all this newness, serious worries arose about the epistemic integrity of the thinkings-up of mathematical structures that had had no prior presence in mathematical practice. The worries were both alethic and epistemic. One wants one's theorems to be true, and to be known to be so without countervailing risk of defeat.[11] But it is a steadfast fact about the human cognitive economy that beings like us make errors, lots and lots of them. Sometimes our proofs are predicated on inconsistent axioms. Sometimes the axioms are fine but the proof rules are leaky. And sometimes key concepts are sloppily defined. We shall shortly see that these three vulnerabilities lie at the motivational core of Frege's foundational mathematics and also for a while of Russell's. They are important enough to have a name. Let's call them the Terrible Three–*false axioms, leaky proofs* and *slovenly definitions*. I'll come back to this unappealing trio in the section to follow.

In the context of the new, the unintuitive and the unproved, these are obvious worries. It is only natural that there would be measures designed to mitigate the risk of mathematical error. In this company, two broad "schools" of risk-tolerance would dominate. On the "conservative" or Weierstrassian side, the toleration of risk was low. On the "liberal" or Riemannian side it was higher. A related wrinkle in this controversy is the importance accorded to the definition of terms prior to their release into the flow of mathematical practice. Weierstrassians favoured precise and explicit definition as a condition of an idea's employment. Riemannians thought that profitable future usage would in time amount to its implicit definition.

It hardly needs saying that the Weierstrassian-Riemann divide is itself neither mathematically well-defined nor in all respects mutually exclusive. It is enough to say that the conservatives carried a low toleration for the risk of error, and liberals were more disposed to invest more heavily in the *post facto* error-correction side. Seen the latter way, mathematics is an experimental science, and a fallibilist one to boot.

2.2 Pre-paradox Logicism

Russell is said to be the co-founder, with Frege, of logicism. This in fact is untrue. Russell was too young to have been its founder.[12] Russell (1872–1970) was twenty-fours younger than Frege (1848–1925) and, in the year that Russell was Seventh Wrangler in Cambridge's mathematics tripos, Frege had brought forth volume I of his *chef d'oeuvre* the *Grundgesetze*.[13] When, nine years earlier, Frege published

[11]Since most epistemologies require that truth be a condition on knowledge, henceforth I'll frame the worry as epistemic.

[12]Notwithstanding some occasional slippage. For example, he endorsed the idea that the functions of which the True and the False were their respective values could be left to arbitrary stipulation.

[13]Frege [9].

the *Grundlagen* in 1884, Russell was a precocious boy of twelve.[14] Russell's principal mathematical influence at Cambridge was Whitehead, and it was Moore who mattered most philosophically. When it came to mathematics, young Russell tried to conform to Whitehead's insistence on mathematical precision and rigour. When it came to the *philosophy* of mathematics he set his cap for the certainties afforded by the philosophical analysis of complex concepts. Russell learned from Moore that some concepts are simple and philosophically undefinable. Though unanalyzable, they were perfectly in order as they stood, and graspable with intuitive certainty.

The principal source of Russell's early logicism wasn't Frege. It was Peano, whom Russell had met at the 1900 International Congress of Philosophy in Paris. In *Principles,* Russell reports that he had formed his logicist convictions and worked out most of the key definitions before he became aware of Frege's logicism. (xviii) While Peano was the inspiration, Russell would soon see that Frege was a considerable improvement on him. So we must stay awhile with Frege.

It only stands to reason that during the heyday of the Weierstrassian-Riemanian turbulence, there would be spasms of anxiety about the epistemic *bona fides* of mathematics' upper reaches.[15] The new mathematics differed markedly from the old. It was a considerably more abstract discipline and its new-fangled concepts were unfamiliar to the point of counterintuitiveness. It made perfect sense to wonder about the intelligibility of transcendental numbers, and it is not at all surprising that ways of relief and reassurance would be sought. A case in point—indeed a signature feature of nineteenth century mathematics—was the *arithmeticization* of analysis, which aimed to define all concepts of analysis in number theory. Speaking to the International Congress of Mathematicians in Paris in 1900, Poincaré summed up the then-present state of analysis:

> *Arithmeticism:* "Today in analysis there remain only natural numbers or finite or infinite systems of natural numbers …. Mathematics, as one says, has been arithmetized."[16]

The arithmeticization of the upper reaches of mathematics was epistemologically motivated. It was thought that arithmetic was the paragon of mathematical truth and the gold standard of utter certainty. It was further supposed that these undoubted virtues would be conferred upon higher mathematics by the ties that bound it to arithmetic.

Then something happened to break the spell. Here is Dedekind writing in 1872, where he reports that when some years earlier he started teaching the differential calculus he

> felt more keenly than ever before the lack of a really scientific foundation for *arithmetic*.[17]

There are two points to take from this. One is that arithmetic was, over time, an intensifying source of worry for Dedekind, and not remotely a merely nagging one.

---

[14]Frege [8].

[15]Gray [15].

[16]Poincaré [30].

[17]Dedekind [5]. Reference here to the Foreword of the first edition.

The other is that it was a long-held worry, considerably pre-dating the opening bars of Frege's *Grundlagen* of 1884:

> … is it not a scandal that our science should be so unclear about the first and foremost among its objects [i.e., the number one], and one which is apparently so simple? Small hope, then, that we shall be able to say what a number is. If a concept fundamental to a mighty science gives rise to difficulties, then it is surely an imperative task to investigate it more closely until those difficulties are overcome. (p. xiv)

We take these expostulations as evidence of the blowback on the tried and true of the epistemic anxieties that had so riled the Weierstrassian tiers of the higher mathematics. I lack the space to go on at length about this most curious reversal of confidence. It will be enough to say for present purposes that what Dedekind, Frege and others felt the want of can be summed up this way:

> *The Terrible Three test*: A mathematical theory is epistemically secure only to the extent that its theorems are provable from irrefutable axioms, air-tight definitions of its basic concepts, and leak-proof rules of demonstration. That is, it must wrestle the Terrible Three to the ground.

In 1888, Dedekind axiomatized number theory. So did Peano a year later, but not less ably.[18] But in short order it was clear that Frege had not found in them the solace he sought. We are left to conclude that the reason why is that Frege thought that Dedekind hadn't managed to pass the Terrible Three test.

In a move that would bear his name evermore, Frege added a codicil to the Terrible Three test, according to which it would be a passable test only if the axioms, definitions and proof rules that validated arithmetic arose in *logic*. Thus was born the doctrine which Frege never called "logicism". The name was assigned it by Abraham Fraenkel in 1928, three years after Frege's death. We should also note that Fraenkel meant "logicism" as a characterization of the position of Whitehead and Russell.[19] As we shall shortly see, logicism in the Whitehead-Russell sense is about the last thing that Frege's doctrine could have been. Frege's proferred rescue-theory for arithmetic was a second-order functional calculus of his own making, upon which he bestowed the name of logic. Russell's would be his and Whitehead's *Principia*,[20] also carrying the name of logic, which differed from *Grundgesetze* in a number of ways, not least in its intended scope. It was designed to accommodate all of mathematics, including a chunk of geometry, not just number theory.

Russell had enough familiarity with Frege's work to have alerted him in 1902 of the paradox occasioned by Basic Law V. Until its discovery, Russell had been a steadfast supporter of the idea that logic was the only epistemically safe harbour for mathematics. Pre-paradox Russell, like Frege, was a thoroughgoing *epistemic* logicist, a quester for mathematical certainty. Post-paradox Russell would be a different kettle of fish.

---

[18] Dedekind [6] and Peano [24].

[19] Grattan-Guinness [14].

[20] Whitehead and Russell [43].

It is generally accepted that the paradox of the set that is and is not its own member marked the death of Frege's logicism. In fact there was another and more basic point of vulnerability. Analyticity was Frege's anti-Terrible Three elixir. His was not the Kantian notion according to which an analytic truth was one in which all information carried by a statement's predicate is already present in its subject. Kant's conception presented Frege with two difficulties. One was that Frege had abandoned the subject-predicate characterization of statements distinction for the newer one between arguments and functions. The other was that Frege did not want the epistemically secure theorems of the *Grundgesetze* to be uninformative in the way that "All bachelors are unmarried men" is. In the *Grundlagen* (p. 4), Frege says that to ascertain whether a proposition is analytic or synthetic, it is necessary that we examine its proof and determine whether it flows incorrigibly from *primitive truths,* that is, from principles that are neither needful nor susceptible of independent proof. If, in turn, primitive truths are wholly general laws which validate the proofs' definitions, then the statement in question is analytic.[21] Frege accepts that analytic truths close the world. Their truth is invariant under all possible permutations of fact. Frege thought that if the problem of the Terrible Three were to be solved, axioms would have to be analytic, definitions analytically tight and proof-rules analytically truth-preserving. But it is easy to see that, on his own telling, Frege cannot pass this test.

> *Frege's Bind*: A proposition is analytic only if it lies in the demonstrative closure of its governing axioms or primitive truths. Since primitive truths are indemonstrable, they are excluded from membership in any demonstrative closure. So no primitive truth is analytic, and the Fregean bailout cannot meet the terms laid out for it.

The question that now presses is how Frege could have missed the bind that he was in. The answer, in large part, was because of another bind that he didn't yet know he was in. Frege had an old-fashioned conception of axioms, virtually the same as Aristotle's notion of first principles. An axiom was a statement that carried the Good Housekeeping Seal of Epistemic Approval. Its epistemic assurance was intrinsic. It was unneedful and insusceptible of proof and fully sealed-off from refutation and falsification. Today, however, "Frege's view is a dinosaur", such is the measure of its old-fashionedness.[22] Frege was caught in a kind of time-warp. On the one hand, he aspired to be a strict Terrible Three-free logicist, which denied him release from Russell's paradox. Thanks to the paradox, it lies with the very idea that the concept of set is empty. Frege's old-fashioned view of axiomhood made matters worse.[23] Even if he were to agree that the paradox made Basic Law V somehow "inoperable"—i.e. unusable as a premiss - he couldn't in all consistency allow that Russell had shown it to be false. How could it have done? Basic laws are beyond refutation. This is a further bind that Frege was in.

---

[21] In about these same words, Frege characterizes the *à priori*, thus collapsing the Kantian distinction between the alethic property of analyticity and the epistemic property of apriority.

[22] Blanchette [2, p. 31].

[23] For a fuller discussion of this difficulty, readers could consult Woods [47].

> *Frege's Double Bind*: Set theory cannot be saved in any of the ways under contemplation at the time, and Frege's own set theory is beyond repair. Conclusion: Set theory is mathematically impossible, and the epistemic certainty of arithmetic cannot be secured by logicism. Indeed it cannot be secured at all.

Russell also saw the paradox as a quite general disaster. The year after his letter to Frege, Russell concedes that "I have failed to perceive any concept fulfilling the conditions requisite for the notion of *class*".[24] This throws as much of a monkey-wrench into Russell's project as it did to Frege's. Before the blow fell, Russell's own method

> will therefore be one of *analysis*, and our problem may be called *philosophical* - in the sense, that is to say, that we seek to pass from the complex to the simple, from the demonstrable to its indemonstrable premises. (p. 2; emphasis mine)[25]

It is important to note, therefore, that the method of analysis was *itself* a regressive procedure, in which one reasoned *backwards* from the received statements of settled mathematical practice to its unanalyzable primitive truths, in whose demonstrative closure the rest of its truths would reside. From time to time, Russell would call the regressive method the "method of analysis."

Russell knew the jig was up for analytical logicism before he published his 1902 paper, which had given a favourable impression of it. It took him awhile to come to his senses, and in *Principles,* he charts a markedly new course for logicism. He grants that there is *no* concept of set, hence nothing for the logicist to analyse. But Russell allows for *mathematical*, as opposed to philosophical definitions, in a sense that "is widely different from that current among philosophers." (p. 15) He goes on to explain that

> it is necessary to realize that definition, in mathematics, does not mean, as in philosophy, an analysis of the idea to be defined into constituent ideas. This notion, in any case, is only applicable to concepts, whereas in mathematics it is possible to define terms which aren't concepts. (p. 27)

Moreover,

> of the three kinds of definition admitted by Peano – the nominal definition, the definition by postulates, and the definition by abstraction – I recognize only the nominal. (p. 112)[26]

Russell sometimes called his new approach the method of *synthesis*, the method of making up new ideas from bits and pieces of old ones. Henceforth his would be *synthetic logicism.*

---

[24] Russell [38].

[25] In "On the axioms of geometry" of 1899, Russell lists seven simple and indefinable concepts: addition, number, order, equality, less, greater and manifold. See Griffin and Lewis [17].

[26] Russell took a somewhat rigourist approach even to mathematical stipulation. He dismisses definition by abstraction, which "… suffers from a wholly fatal formal effect: it does not show that only one object satisfies the definition." (p. 114).

In 1901–02 Russell accepted that sets were needed for modern mathematics. But he agreed with Frege that, thanks to the paradox, the very concept of set was philosophically empty. So he *stipulated* sets into nominal existence by way of mathematical definition. In so doing, a further difficulty arose.[27] Should the made-up axioms of set-theory turn out not to be true, the theory's valid proofs would become merely conditional. They would establish that if the premisses are true, then the conclusion is also true. Left unresolved, Russell's brand of synthetic logicism would have conditional standing only. This is the so-called if-then-ism problem. It is contested as to whether Russell saw this as a one.[28]

Like the G*rundgesetze, Principia Mathematica* is a work of daunting complexity and technical virtuosity. It is just as it stands, a work of great architectural sublimity. Nothing is easy in *Principia,* and hardly anything is free. It is not until proposition * 110.643 on page 83 of the second volume that we are treated to the proof that 1 + 1 = 2. It would be lovely to linger awhile longer, but it is past time that we return to the disclosures of 1907. Let's close this section with the summary observation that now that there is no concept of set, there is nothing available to Russell for philosophical analysis. This puts Russell in a bind. In the Preface to *Principles,* Russell says that in the absence of conceptual analysis, philosophy of mathematics is impossible.[29] And with that impossibility comes another. It is impossible to provide epistemic surety for the theorems of mathematics.

> *Russell's Bind*: Since classes are essential for logicism and yet are beyond the reach of conceptual analysis, logicism lies beyond the power of philosophy to deliver.

What is more,

> *Russell's Double Bind:* Even if classes *were* philosophically analyzable, its primitive truths would be open to the trouble that plagued Frege's indemonstrables.

Hoist on the same pétard, Frege and Russell parted ways. Frege gave up on logicism as an irretrievably lost cause. Russell changed the subject but not the name. From that point onwards his logicism would be the stipulated product of mathematical definition. It would be the logicism effected by synthesis.

2.3 Russell's Post-paradox Logicism

On March 9, 1907, Russell read to the Cambridge Mathematical Club an essay entitled "The regressive method of discovering the premises of mathematics," a paper that wasn't published until after his death sixty-three years later. We owe its exposure to Douglas Lackey in [41] and more recently and, in a more directly focused way, to Andrew Irvine.[30] In the year that Russell read the regression paper

---

[27] See, for example, Putnam [31], Griffin [16] and Krall [21].

[28] For reservations, see Griffin [16].

[29] *Principles*, xviii.

[30] Irvine [20].

he published a paper entitled "The study of mathematics" which spoke approvingly of analytic logicism.[31] Russell appears to have written this paper in 1902, the same year of his letter to Frege announcing the contradiction that flowed from Basic Law V of the *Grundgesetze*.[32] In fact, Russell had come upon the contradiction in the spring of 1901.[33] This was the year in which Russell also published "Recent work on the principles of mathematics".[34] Russell's next post-paradox publication was *The Principles of Mathematics* of 1903. The interval from Spring 1901 to December 1902, when Russell completed *Principles* reflect a *volte face* in Russell's philosophy of mathematics, a radical departure from his own former philosophical self.

Here in brief summary form is the state that logicism was in in 1907 and counting. Axiomatization was a recognized regressive enterprise. Frege's own *Grundgesetze* had been regressively structured. Pre-paradox Russell reserved pride of place for axiomatizations effected by conceptual analysis, which would expose the axioms as primitive truths whose certainty is intuitively assured. After the paradox, Russell gave up on conceptual analysis for logicism, but he did not give up on logicism. He retained regressive measures for axiomatizing mathematics, whose primitive principles would now be advanced by mathematical stipulation. Since mathematical definitions aren't certainty-producing, the best we can hope for is to find reason for believing one's axioms to be true. For this to work, grounds would have to be found for believing to be true axioms which hardly appear so—notably the axioms of infinity, choice and reducibility. If successful, axioms needn't be intuitive or carry the look of truth or even plausibility to qualify for reasoned believability. To this end, Russell frames a question along the following lines; "What are the unproved propositions $p_1, \ldots, p_n$ in whose demonstrative closure the known or accepted truths of mathematics are wholly included." The follow-up question asks for the methods by which these propositions are to be found.

In contexts of regressivist enactment, Russell draws an odd-seeming distinction between empirical propositions and logical propositions. An empirical proposition is an obvious truth of mathematics such as the proposition that $1 + 1 = 2$. A logical proposition is one that implies an empirical proposition of this kind. Generalizing, in the first instance Russell is looking for the least class of logical propositions in whose demonstrative closure all *obvious* mathematical truths are contained. From this, Russell conjectures that if C+ is got by addition of further logical propositions to C and that all of mathematics, both the obvious and the nonobvious, can be shown to lie in *its* demonstrative closure, that would be reason to believe that the propositions of C+ are true. Whatever its details, the regressive method is a *premiss*-search, not a search for newly revealed consequences. It is not, as we shall see, a hypothesis-selection search.

---

[31] Russell [36, 29–44]. Reprinted in the first edition of Russell (1910).

[32] Bertrand Russell, Letter to Frege, in van Heijenoort [39, 40]. Frege's letter in reply follows at pages 127–128.

[33] Russell appears to have been unaware of Zermelo's derivation of it the year before. See Hallett [19].

[34] Russell [35]. Reprinted with revisions and a new title in Russell (1918).

This brings us to what's distinctive about Russell's 1907 paper. It is not made so by its endorsement of regressive methods. It was made so by the way in which it answered the *follow-up* question. The way to determine whether a proposition qualifies as a p_i is by trial and error, i.e., by examining how much of received mathematics resides in its demonstrative closure. In so thinking, Russell places himself in the descendant-class of his godfather, J. S. Mill, whose *A System of Logic* (1843) sees logic as a science on par with the sciences of nature, whose basic laws are arrived at *inductively.* In a general sort of way, this shows Russell in a now more Riemannian frame of mind from that of his earlier leanings towards conceptual analysis. More to the point, though he took no notice of it, is that Russell has aligned himself to the inductive measures advocated by logic's very founder, whereby the practitioner of a science–mathematics and logic included—are seized of the binding authority of its first principles.[35] As pointed out in Irvine [20], Russell has framed his case for the reasonable believability of the axioms of *Principia* on the model of how the believable laws of the physical sciences are arrived at. I find Irvine's case to be a convincing one, and the case he attributes to Russell to have genuine merit. But the matter before us is not whether Russell's reasoning in 1907 was sound, but rather whether it was abductive.

3 Abduction and Regression

3.1 Curtain

The idea that regressivism is abductive is said to have been suggested by Scott Kleiner in conversation with Irvine.[36] Irvine reports the suggestion without further comment, beyond saying that "the regressive method is similar to Peirce's abduction" (*idem.*) Perhaps that is so. The question is whether it was similar enough to matter for either enterprise. My answer to that question is in the negative. There are too many of the Peircian boxes that Russell's project doesn't tick. Let's begin with

Russell's Question: What would it take to make it the case that the epistemically uncertain propositions stipulated in *Principia* are propositions we have reason to believe to be true?[37]

---

[35]This view is developed in Aristotle's *Posterior Analytics* and *Metaphysics* [1]. Details can be found in Woods [47].

[36]Irvine [20] reports that suggestion in ft. 26, p. 322. Kleiner is a philosopher of science at the University of Georgia.

[37]Of course, volume one of *Principia* has yet to appear. Shouldn't we instead take Russell to be thinking of the logic of the *Principles*? I think not. On May 13, 1903, a few days after its publication, Russell had lost confidence in *Principles*, writing that "it seems to me a foolish book, and I am ashamed to think that I have spent the best part of six years upon it. Now that it is done, I can allow myself to believe that it was not worth doing—an odd luxury!" See Griffin [18]. Even allowing for Russell's habit of self-effacement, it is better to locate his logical thinking in 1907 in Russell [37] in van Heijenoort [10], pages 152–182, which would appear a year later and would be absorbed into *Principia,* whose first appearance was two years after that.

Of course, that is not a question answerable by a Peircian abduction. *No* question that triggers a Peircian abduction is answered by its successful enactment. The best that it could deliver is that, while there is no reason to believe the propositions of *Principia* to be true, there is reason to *suspect* they are true. However, reasoned suspicion of the truth is not what "The regressive method" is after. It undersells the Russellian line.

Let's turn now to the selection step. What should we take H to be? Suppose we put it that H is the hypothesis that mathematics is derivable from *Principia.* Slot this H into the antecedent of the subjunctive conditional at line 10 of the Gabbay-Woods schema. This would give "If adding to what we already know, the proposition that mathematics is derivable from *Principia*" resulted in a true statement, then Russell's question would be answered." Clearly, however, it does not answer Russell's question. The selection-step is misplayed. What is more, even had that step not been misplayed, the inference-step would be misplayed by any sentence, other than "Hence there is reason to *suspect* that mathematics is derivable from *Principia*". But, again, that would undershoot Russell's target. Indeed, it would be the last thing from which Russell would derive cause to submit to Peircian stress-test. Let's be clear about this. Russell does indeed want to submit *Principia* to a stress-test. But his decision misnegotiates the Peircian action-step. Peircian action-steps are always occasioned by suspicions of something's truth in the absence of any reason to believe to be so. That was never Russell's position with respect to *Principia.* So, as I now propose, all this casts serious doubt on the abductive regression thesis. It misnegotiates too many steps in the abductive inference schema for serious belief.

References

1. Aristotle: The Complete Works of Aristotle: The Revised English Translation, vol. 2. In: Barnes, J. (ed.) Princeton University Press, Princeton (1984)
2. Blanchette P.A.: Axioms in Frege. In: Rossberg, M., Ebert, P. (eds.) Essays on Frege's Basic Laws of Arithmetic. Oxford University Press, Oxford
3. Chiffi, D., Pietarinen, A.-V.: Abductive inference with a pragmatic framework. Synthese. 1007/s11239-018-1824-6
4. Chiffi, D., Pietarinen, A.-V.: The extended Gabbay-Woods Schema and scientific practice. In: Gabbay, D.M., Magnani, L., Park, W., Pietarinen, A.-V. (eds.) Natural Arguments, A Volume in the Tributes Series. College Publications, London (2019)
5. Dedekind, R.: Essays on the Theory of Numbers, Dover, New York. Reference here to the foreword of the first edition. First published in 1872 (1963a)
6. Dedekind, R. The Nature and Meaning of Numbers. Reissued in two volumes by Dover. First published in 1888 (1963b)
7. Ferreirós, J., Gray, J. (eds.): The Architecture of Modern Mathematics: Essays in History and Philosophy. Oxford University Press, New York (2006)
8. Frege, G.: Die Grundlagen der Arithmetik. Wilhelm Koebner, Breslau (1884). Translated into English as Foundations of Arithmetic, J. L. Austin, Blackwell, Oxford (1950)
9. Frege, G.: Grundgesteze der Arithmetik. Herman Pohle, Jena (1893/1903). In: Ebert, P.A., Rossberg, M., Wright, C. (eds.) Translated into English as Basic Laws of Arithmetic. Oxford University Press, Oxford (2013)

10. Frege, G., Letter to Russell. In: van Heijenoort, J. (ed.) From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931, pp. 127–128 (1967)
11. Gabbay, D.M., Woods, J.: The Reach of Abduction: Insight and Trial. North-Holland, Amsterdam (2005)
12. Goldfarb, W.: Logtic in the twenties: the nature of the quantifier. J. Symbol. Logic **44**, 351–368 (1979)
13. Goldman, A.: A causal theory of knowing. J. Philos. **64**, 357–372 (1967)
14. Grattan-Guinness, I.: The Search for Mathematical Roots, 1870–1940. Princeton University Press, Princeton (2000)
15. Gray, J.J.: Anxiety and abstraction in nineteenth-century mathematics. Sci. Context **17**, 23–47 (2004)
16. Griffin, N.: New work on Russell's early philosophy. Russell **2**, 69–83 (1982)
17. Griffin, N., Lewis, A.C. (eds.) The Collected Papers of Bertrand Russell Volume II Philosophical Papers, 1876–1899, pp. 390–415. George Allen & Unwin, London (1982)
18. Griffin, N. (ed.) The selected letters of Bertrand Russell, vol. 1. In: The Private Years, 1884–1914, Letter 120. Houghton Mifflin, Boston (1992)
19. Hallett, M.: Cantorian Set Theory and Limitations of Size. Clarendon Press, Oxford (1984)
20. Irvine, A.D.: Epistemic logicism and Russell's regressive method. Philos. Stud. **55**, 303–327 (1989)
21. Krall, A.: The aim of Russell's logicism. Synthese **191**, 1493–1510 (2014)
22. Magnani, L.: The Abductive Structure of Scientific Creativity: An Essay on the Ecology of Cognition, Cham. Springer, Switzerland (2017)
23. Park, W.: Abduction in Context: The Conjectural Dynamics of Scientific Reasoning, Cham. Springer, Switzerland (2017)
24. Peano, G.: The principles of arithmetic, presented by a new method. In: van Heijenoort, pp. 85–97. First published in 1889 (1967)
25. Peirce, C.S.: Description of a notation for the logic of relatives, resulting from an amplification of the conceptions of Boole's calculus of logic. Memoirs Am. Acad. **9**, 317–378 (1870)
26. Peirce, C.S.: The logic of relatives. In: Peirce, C. S. (ed.) Studies in Logic, pp. 1870-203. Little, Brown & Co., Boston (1883a)
27. Peirce, C.S.: Second intentional logic. In: Peirce, pp. 56–58 (1883b)
28. Peirce, C.S.: Reasoning and the Logic of Things: The Cambridge Conference Lectures of 1898. In: Kettner, K. L. (ed.) With an Introduction by Kettner and Hilary Putnam. Harvard University Press, Cambridge, MA (1992)
29. Peirce, C.S.: Collected Papers. In: Hartshorne, C., Weiss, P. (vol. I–VI) Burks, A.W. (vols. VII and VIII) (eds.) Harvard University Press, Cambridge, MA (1931–1958)
30. Poincaré, H.: Du rôle de l'intuition de la logique in mathematiques. Compte Rendue der Deuxième Congrés International des Mathématiciens, Paris (1902)
31. Putnam, H.: The thesis that mathematics is logic. In: Shoenman, R. (ed.) Bertrand Russell: Philosopher of the Century, pp. 273–303. Allen & Unwin, London (1967)
32. Putnam, H.: Peirce as a logician. Hist. Math. **9**, 290–301 (1982)
33. Quine, W.V.: Peirce's Logic. In: Quine [34]
34. Quine, W.V.: Selected Logic Papers, pp. 258–265. Harvard University Press, Cambridge, MA (1995)
35. Russell, B.: Recent work on the principles of mathematics. Int. Month. **4**, 83–101 (1901). Reprinted with revisions and a new title in Russell (1918)
36. Russell, B.: The study of mathematics. New Q. **I**, 29–44 (1907). Reprinted in the first edition of Russell (1910) Bertrand Russell Philosophical Essays. Longmans, Green & Co., London (1910)
37. Russell, B.: Mysticism and Logic. Longmans Green, London (1918)
38. Russell, B.: The Principles of Mathematics. Cambridge University Press, Cambridge (1903); second edition. Allen & Unwin, London (1937)
39. Russell, B.: Letter to Frege. In: van Heijenoort (ed.), pp. 124–125 (1967a)

40. Russell, B.: Mathematic logic as based on the theory of types. In: van Heijenoort (ed.), pp. 152–182 (1967b)
41. Russell, B.: The regressive method of discovering the premises of mathematics. In: Lackey, D., Russell (eds.) Essays in Analysis. George Allen & Unwin, London (1973)
42. Tappenden, J.: Metatheory and mathematical practice in Frege. Philos. Topics **25**, 213–264 (1997)
43. Whitehead, A.N., Russell, B.: Principia Mathematica, vol. 3. Cambridge University Press, Cambridge (1910–1913). Second edition, 1925 and 1927
44. Woods, J. Errors of Reasoning: Naturalizing the Logic of Inference, vol. 45. Studies in Logic. College Publications, London (2013); reprinted with corrections 2014
45. Woods, J.: Reorienting the logic of abduction. In: Magnani, L., Bertolotti, T. (eds.) Springer Handbook of Model-Based Science, pp. 137–150. Springer, Berlin (2017)
46. Woods, J.: Truth in Fiction: Rethinking its Logic, vol. 391. Synthese Library: Springer, Cham, Switzerland (2018)
47. Woods, J.: What did Frege take Russell to have proved? Synthese. https://doi.org/10.1007/s11229-019-02324-4 (2019)

Abduction, Mathematics, and Logic in Creative Discovery

The Place of Logic in Creative Reason

Atocha Aliseda

Abstract In this text, I put forward the thesis that creativity and logic do not exclude each other. I depart from the characterization of a creative product as that which is novel and useful. I presuppose there is some kind of method for generating that new product, and in this respect, I rely on Peirce's formulation of abduction. A central point to be discussed concerns whether Peirces' abduction may be considered as a logic of synthetic reasoning; to what extent its products are genuinely novel. In my defense of above's thesis, I shall restrict to new scientific hypotheses generated from previous concepts. I consider that this modest analysis will nevertheless shed some light on the scope of this thesis.

1 Introduction: Creative Reason

Creative Reason has a place in the process of invention when we conceive of an idea for the first time or discover a new scientific theory. It is, however, necessary to establish a position with regard to the following question: What do we understand by creativity? The lack of even minimal agreement on the meaning of *creativity* as a concept is notorious. For the purpose of this work I shall adopt, although tentatively, a characterization proposed in the field of cognitive psychology (see Sternberg and Lubart 1999: 3) which defines a creative product as one that is both *novel* (original,

Research for this article was supported by the project Logics of Discovery, Heuristics and Creativity in the Sciences granted by the National Autonomous Univeristy of Mexico. A previous version of this text was publisehd in Aliseda [3].

A. Aliseda (✉)
Instituto de Investigaciones Filosóficas, National Autonomous University of Mexico, México CDMX 04510, Mexico
e-mail: atocha@filosoficas.unam.mx

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_7

unexpected) and *useful* (appropriate, adaptable). On the other hand, when I refer to Creative Reason, I am simply assuming that there is a certain method involved.[1]

The questions which interest me in this text regarding the origin of Creative Reason and its justification are related to inquiries into *synthetic reasoning*, which go back to Immanuel Kant and were reformulated by Charles S. Peirce, as the following passage testifies:

> According to Kant, the central question of philosophy is 'How are synthetical judgments a priori possible?' But antecedently to this comes the question how synthetical judgments in general, and still more generally, how synthetical reasoning is possible at all. When the answer to the general problem has been obtained, the particular one will be comparatively simple. This is the lock upon the door of philosophy. (CP [25]:5.348, cited in Hookway [17]: 18)

Peirce's analysis of *ampliative reasoning* or, as he preferred to call it, synthetic reasoning has two fundamental aims: on the one hand, to justify its possibility; on the other, to develop a method for putting it into practice. Pierce found the answer to the first in his metaphysics; the second is a matter addressed by some contemporary formal research in logic, above all that referring to *abductive logic*.[2]

The following question captures very well the place of Logic in Creative Reason: *Is there a logic of scientific discovery*? While this question has its origin in Ancient Greece and traverses much of the history of philosophy, it is worth stressing its place in the context of contemporary philosophy of science and in its later adoption by cognitive psychology. In the second part of this article, I shall go into this question in depth, drawing on the work of Peirce, who conferred a logical formulation upon abduction, and which I shall review in detail in relation to the question of whether this is effectively a logic of synthetic reasoning and in what sense its products are *novel*.

In this text I shall argue, in particular, for the thesis that *logic and creativity are not mutually exclusive*. It is worth clarifying that I shall be analyzing this thesis in a somewhat restrictive way. As regards creativity, I shall be concerned only with the generation of new scientific hypotheses and, as far as Logic is concerned, I shall limit myself to the formal interpretation that has been given to Peirce's formulation. I believe that this modest analysis will enable us to appreciate both the scope and the limitations of this thesis.

---

[1] There are, of course, other reasons identified in Philosophy and related disciplines: Kant's Pure Reason, the Historical Reason of Ortega y Gasset and, of course, the Poetical Reason of María Zambrano.

[2] See Hookway [17] for an analysis of the metaphysical aspect of ampliative reasoning in Peirce. For an argument in favor of the utility of ampliative logics for modeling creative processes involved in the genesis of new ideas and the generation of scientific theories, see Aliseda [4]; for a panorama of recent developments in abductive logic, see Aliseda [2].

2 Logics of Scientific Discovery

2.1 Contexts of Research in the Philosophy of Science

The question regarding the types of logic involved in scientific discovery is relevant in the framework of the methodological distinction proposed by Hans Reichenbach between the *context of discovery* and the *context of justification* [30]. While the former has to do with the origin and genesis of scientific theories, the latter entails the epistemic and methodological questions *proper* to the philosophy of science, those that support the truth and reliability of scientific theories. As regards the relation between logic and creativity, Reichenbach makes his position very clear when he states that "the act of discovery escapes logical analysis". (Reichenbach [30]: 231).

This distinction of contexts is one of the pillars on which "the received view" has been built up in the philosophy of science; this concerns itself with the context of justification and has to do exclusively with the so-called "finished research report" of a theory. This certainly leaves a wide space for the context of discovery in the research process.

Nonetheless, many authors have regarded this distinction as both insufficient and unsatisfactory and have put forward proposals of their own. Some identify an intermediate stage between the conception and the justification of a new theory. Claude Savary, for example, characterizes this stage as one of "working with ideas" (Savary [32]); Larry Laudan speaks of a "*context of pursuit*", a sort of "*nether region*" between the initial and final contexts (Laudan [22]: 174); he also introduces another dimension into the study of the context of discovery which consists in distinguishing between a narrow point of view and another broader one. For Savary, the problems of discovery are restricted to the initial conception of an idea, such as the famous "Eureka moment"; while for Laudan, they are implicated in the whole process setting off from the conception of a new idea until the positing of an idea subject to final justification.

Another way of proceeding is to extend the frontiers of the context of justification in order to deal also with the matter of evaluation, in particular when the truth of the theory is not the only thing that interests us. One consequence of this vision is the proposal of Theo Kuipers, who re-named the context of justification *context of evaluation* [19]. Yet another proposal is that of Alan Musgrave, who prefers the term *context of appreciation* (Musgrave [23]:20). In turn, the latter renames the context of discovery *context of invention*, in order to avoid the apparent contradiction that arises when we talk of the discovery of a hypothesis, since "discovery" is a word that suggests a successful process, presupposing the truth of what is discovered.

On the other hand, when the emphasis is on scientific practice, the proposal of Javier Echeverría [11] must not be passed over. He makes a clear distinction between four contexts: "education", "research", "application" and "evaluation". This point of view is attractive because it does not oblige us to situate creativity or discovery in a single context; both can find a place in any of the contexts. Consequently, not only can the original distinction between those of discovery and justification be further subdivided, but it also becomes clear that their frontiers are not so clearly defined. A question that emerges naturally is whether the context of discovery—or any other context one wishes to consider—is susceptible to philosophical reflection and whether it lends itself to logical analysis. The latter point is precisely what we are concerned with here.

Questions regarding scientific discovery have come to the fore in contemporary philosophy of science, although "most often with the intention of excluding the matter from epistemological analysis" (Pérez Ransanz [28]:347), and Karl Popper pronounced an—apparently categorical—disqualification. The common idea regarding the position of this author is that problems of discovery cannot be studied within the frontiers of methodology, for he explicitly denies the possibility of a logical reconstruction of the processes of discovery, and considers that its study is a matter that pertains to psychology. This position is clearly stated in the following often quoted passage:

> … I shall distinguish sharply between the process of conceiving a new idea, and the methods and results of examining it logically. As to the task of the logic of knowledge—in contradistinction to the psychology of knowledge—I shall proceed on the assumption that it consists solely in investigating the methods employed in those systematic tests to which every new idea must be subjected if it is to be seriously entertained.
>
> Some might object that it would be more to the purpose to regard it as the business of epistemology to produce what has been called a 'rational reconstruction' of the steps that have led the scientist to a discovery—to the finding of some new truth. […]. But this reconstruction would not describe these processes as they actually happened: it can give only a logical skeleton of the procedure of testing. Still, this is perhaps all that is meant by those who speak of a 'rational reconstruction' of the ways in which we gain knowledge.
>
> It so happens that my arguments in this book are quite independent of this problem. However, my view of the matter, for what it is worth, is that there is no such thing as a logical method of having new ideas, or a logical reconstruction of this process. (Popper [27]: 8–9)

A response to Popper's position has been offered by one of his most prominent—and at the same time critical—students, Imre Lakatos, who states, firstly "There is no *infallibilist* logic of scientific discovery leading infallibly to results" and, secondly, that "Popper, who has laid the basis of this logic of discovery, was not interested in the meta-question of what is the nature of this investigation, so he did not realize that it is neither psychology nor logic, but an independent field, the logic of discovery, **heuristics**. (Lakatos [20]: 167, my emphasis)

In Lakatos' vision, heuristics is that method of scientific discovery half-way between logic and psychology. However, as I have argued elsewhere [5], Popper's logic of scientific inquiry points in the direction of some fundamental mechanisms that fall under the heading of the study of discovery (understood in a broad sense) and also has points of coincidence with the dominant current of the so-called "friends of discovery", headed by Herbert Simon and which I shall consider in the following section.[3]

[3] See Pérez Ransanz [28] for a more detailed analysis of Popper's position on his "two notions of discovery".

2.2 Heuristics in Cognitive Psychology

The subject of scientific discovery has found a privileged position beyond the philosophy of science in cognitive science. I shall mention briefly the positions of two of its pioneers, Herbert Simon and Paul Thagard.

In his essay "Does Scientific Discovery Have a Logic?" Simon sets out to refute Popper's general argument, re-construed for his purposes as follows: "If 'there is no such thing as a logical method of having new ideas', then there is no such thing as a logical method of having small new ideas" (Simon [34]:327).

The ambitious project of a logic of scientific discovery, one that reveals in all its complexity the process I am trying to capture, becomes, in Simon's hands, a much less ostentatious process. Before detailing his proposal, he tells us (in reference to the examples he cites): "*Their modesty as instances of discovery will be compensated by their transparency in revealing underlying process.*" (Ibid.)

This approach, as brilliant as it is humble, enables Simon both to establish the types of problems that are to be analyzed and to specify the methods to be used. For Simon and his followers, *scientific discovery is a problem-solving activity*.[4] Although there is no effective method by means of which to achieve scientific discovery as a problem-solving activity, it can be characterized by way of strategies. The key concept in all this is heuristics. Heuristic methods of discovery are characterized by selective searches with fallible results. That is to say, although they do not guarantee success in arriving at a solution, the search in the problem space is not blind, but selective according to a predefined strategy.[5]

Popper's and Simon's postulations have both convergences and differences. Both subscribe to the vision of science as a problem-solving activity, though differ on just what they consider to be the *Logic* of scientific discovery. While for Popper ideas are generated by the method of *Conjectures and Refutations*, which can be interpreted

[4]If this slogan affirms something that seems obvious nowadays when referring to scientific practice, it is in fact an idea that the methodology of science imported from cognitive psychology [1]. When Simon characterizes scientific discovery as a "problem-solving activity", this translates into computational programs with tools of their own for modeling scientific discovery. Such programs exemplify the possibility of automatic simulation of scientific discoveries, oriented towards modeling fundamental mechanisms that point towards conditions for creative inferences sufficient for leading to the generation of new knowledge. In the context of this conception, a distinction is provided between two types of problems—well-structured versus illstructured ones—, and the aim of finding a logic of discovery concentrates on well-structured problems. The latter type is one for which a clearly defined criterion of testing exists and for which there is, at least, a problem space in which the initial state and the final state can be represented, and where all the remaining intermediate states can be reached with the appropriate transitions between them. An ill-structured problem lacks at least one of the above conditions.

[5]The authors also distinguish between *weak* and *strong* methods of discovery. The former comprise the type of problem solving used in novel fields; they are characterized by their generality, since they do not demand deep knowledge of the particular field. In contrast, the strong methods are used for cases in which our knowledge of the field is extensive, and are especially designed for a specific structure. The weak methods include heuristic methods of generation and testing, and analysis of ends and means, in order to build up explanations and solutions for given problems. For details of the approach of Simon and his followers, see Langley et al. [21].

as a method of 'blind search', Simon and his team develop a theory for upholding the position that ideas are generated by means of a method of 'selective search'. The latter approach enables the way in which *new* ideas can be generated to be captured as such, and concerns heuristics.[6]

In the same spirit of the "friends of discovery", Paul Thagard [37] proposes a new field of study, namely the *computational philosophy of science*, an approach that brings together psychology and the history and philosophy of science, all with a view to scientific discovery and with a particular emphasis on cognitive patterns and key notions in the philosophy of science relevant to this matter: explanation, confirmation, falsification, evaluation, induction, abduction and revision of theories.

The vision of scientific discovery as an instance of Simon's problem-solving activity as well as the conception of a *computational* philosophy of science conceive the inquiry into the processes of scientific discovery as part of the methodology of science. The interaction between the different disciplines that make up this study promotes the importation of computational tools into the philosophy of science, with the aim of modeling the dynamics of scientific knowledge, including its generation and development.

3 Abduction: The Logic of Synthetic Reasoning

My interest here is to explore in greater depth the question of the logics of discovery from the point of view of Logic itself; in this text I shall limit myself to tracing the origins of such logics and analyzing some aspects of them. The first philosopher to propose a logical formulation for abductive reasoning was Charles Sanders Peirce (1839–1914); his influence on those who represented the "received view" was, however, minimal or non-existent. On the other hand, his importance was recognized by certain philosophers who reacted against the prevailing current, precisely because of their concern to investigate matters of relevance to the context of discovery (a prominent case being that of Hanson [15]). Nevertheless, Peirce's notion of abduction is difficult to analyze, since his conception of it evolved throughout his philosophical trajectory, and several different versions appear as his thinking evolved.[7] At the same time, his conception is interlaced with epistemological and semiotic aspects of his philosophy that are equally complex.

For the purpose of this paper, I shall focus on the logical formulation of abduction that emerges from the third and last stage of his thinking, and shall analyze it in the light of the thesis that concerns this text, namely that logic and creativity are not mutually exclusive.

[6]For a deeper analysis comparing these approaches, see Aliseda [5].

[7]For a detailed study, which clearly distinguishes three stages in the evolution of the notion of abduction in Peirce, see Fann [12]. Another key reference is Anderson [7], for an analysis of abduction and creativity.

For Peirce, abductive reasoning is fundamental in every human act of inquiry. Abduction plays a role in perception: "*The abductive suggestion comes to us as a flash*" (CP [25]:181). It also plays a fundamental role in the generation of hypotheses and new ideas in scientific activity: "*Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new ideas*".

As proposed by Anderson [6], abduction seems to be both "an act of intuition and one of inference", which suggests a double aspect: both intuitive and rational. The presence of these two aspects, apparently antagonistic, has invariably caused confusion in scholars of Peirce. In general, only one of these is considered for the analysis of abduction ([18, 33, 39]). Some other critics have interpreted this duality as Peirce's dilemma and concluded that the philosopher lacked a coherent vision of the nature of abduction [13]. There also exists another approach—in addition to the above mentioned one of Anderson—that seeks to give a sense to these two aspects, the intuitive and the rational [8, 24], and this is precisely the line I follow in this text in support of my own thesis: logic and creativity are not mutually exclusive. The abductive formulation is as follows (CP 25:189):

The surprising fact, C, is observed.

But if A were true, C would be a matter of course.

Hence, there is reason to suspect that A is true.

This formulation has been the starting point for many studies in both philosophy and computing. In general its formal interpretation is as follows:

$$\frac{\begin{array}{l}C\\ A \rightarrow C\end{array}}{A}$$

Implicit is the extra-logical specification that the status of the conclusion is *tentative*; otherwise one would incur in the *fallacy of affirming the consequent*. Likewise it is assumed that the second premise is part of the background theory in question.

Aside from its *explanatory capacity*, Peirce considers a further two criteria for a successful hypothesis, namely, *testability* and *economy*. An abductive hypothesis can be regarded as explanatory if it accounts for the facts in accordance with the formulation; its status remains, however, tentative until it is put to the test, thus achieving empirical corroboration and yielding place to the third criterion: that of economy. The latter answers to the practical problem of managing an endless series of hypotheses that satisfy the former criteria.[8]

I shall now analyze some critiques of the Peircean formulation, above all as regards its creative role in respect of the generation of hypotheses. As regards the formal

[8] Some approaches attempt to capture Peirce's economy criterion as a subsequent process for choosing "the best explanation", since there may be many formulae that fit this logical formulation and that are all subject to empirical testing, without being ideal hypotheses. For the interpretation of abduction as *inference to the best explanation*, see Harman [16] and Douven [10]. For a discussion on whether Peirce's notion of abduction is one of inference to the best explanation, see Campos [9].

interpretation of the abductive formulation, we have the following in Frankfurt's words:

> Clearly, if the new idea, or hypothesis, must appear in one of the premises of the abduction, it cannot be the case that it originates as the conclusion of such an inference; it must have been invented before the conclusion was drawn [13]: 594.

This statement points to the *temporal sequence* of the invention of elements that make up the formulation, which in turn directs us to the very meaning of *novelty*. Let us now consider the former. Anderson argues in favor of the simultaneous generation of the two occurrences where "A" appears: "In this sense, *the 'A' in A-C(sic) is logically prior to the 'A' in the conclusion. However, it is also clear that the two 'A's' are one and the same. Therefore, in temporal terms, they may be simultaneously arrived at; it does not follow that A* 'must have been invented before the conclusion was drawn'" [7]:35; and concludes, "… on Peirce's view it is possible for the hypothesis and its abductive application to occur together. Therefore, abductions may be insightful and original and still have logical form." [6]:157. The new, the creative, does not exclude logic; it is part of it.

I shall now examine the idea of *novelty*, as concerns this formulation. What counts as something completely new? According to Frankfurt, the conclusion does not offer anything genuinely new, since it is already present in the premises. But if the creative act consists in identifying something already known A → C (a natural law, for example) and on this basis to suggest A (as the cause of C, for example), then we are, on the one hand, upholding Anderson's argument of simultaneous temporality and, on the other, affirming that the creativity lies precisely in establishing new connections between the elements that present themselves to us, in this case as a surprising fact represented by C and by an associated law (A → C) so as, in consequence, to propose A.

In order to appreciate the functioning of the abductive formulation, let us briefly look at some elements of the reconstruction of a scientific discovery of the eighteenth century attributed to Antoine Lavoisier.[9] This scientist is known for his contributions that were of key importance for the conformation of modern chemistry; here I shall address his research into the nature of combustion. As a product of his observations that certain metals gain in weight as a consequence of combustion, he conjectured that this process entails in some way combination with air, thus refuting the theory of phlogiston. His observations that certain metals effervesced (produced bubbles) during the process of combustion led him to conjecture that such metals "contained air". The latter hypothesis can be represented in the abductive schema as follows:

C	Effervesces (m)
A → C	Contains (x, air) → Effervesces (x)
A	Contains (m, air)

[9] A detailed reconstruction of Lavoisier's discovery as a case of an "abductive rule", can be found in Thagard [38].

While the observations of the increase in weight of metals as a consequence of combustion can be generalized to all metals through an (enumerative) inductive inference, the conjecture that effervescent metals contain air is initially the product of an abductive inference. This schema represents the connection made by Lavoisier between an observation (a certain metal effervesces) and the empirical regularity alluded to (the metals that contain air effervesce), in order to establish precisely a hypothesis (the metal observed contains air), that in turn is generalized in order to conclude that the metals in combustion contain air.

Having analyzed the idea of novelty illustrated in this schema, I shall go on to analyze the scope of the (formalization of the) original formulation of Peirce. In the first place, it could well be argued that it fails to capture in itself either the fact of C being novel, or any of the other criteria proposed by Peirce. Furthermore, note that the interpretation of the second premise has no reason to be the classical material implication; it could be a logical implication of another kind, or even a (computational) process whose input would be A and its output C.

Indeed, today there as many logical as there are computational systems that implement the abductive formulation in somewhat sophisticated ways. For example, in *adaptive logics*, when an abductive conclusion is generated, it is added as a line of the proof with a tentative status. If in later lines new information is added that refutes this hypothesis, a *strategy* enters into operation that marks the line in which it appears (and perhaps other lines that depend on that) and blocks it until it chances to be "resuscitated" as a product of later inferences. Other abductive approaches exist whose aim is to construct algorithms in order to generate abductive hypotheses.[10]

As regards the question whether these formal systems produce genuinely new elements, we must acknowledge that the abductive formulation, even in its most sophisticated forms, reveals an intrinsic characteristic of Creative Reason in its *formal* version: new ideas are always a product of combinations of already existing ideas. This vision is manifested clearly as a "syntactic restriction of language", in the sense that in abductions, *new* elements can only be generated in the framework of the language in question.[11] At first sight this reflects a limitation; at least it underlines the fact that in formal models of creativity there is no place for the generation of new concepts, a phenomenon usually associated with *Revolutionary Science*. In so far as they are creative hypotheses, abductions modeled in formal frameworks refer to facts or relations between them; but when a hypothesis is based on a new theory that involves the creation of a new concept, such models have neither the language nor the mechanisms to generate totally new products. Even so, the generation of new facts and of relations between empirical facts that explain a phenomenon are authentic creative products, inasmuch as they arise from new connections and combinations between existing ideas. No few scientific discoveries have been the product

[10]For an introduction and panorama of the logic of abduction in which three different logical approaches are reviewed (as logical inference, as computational process and as epistemic change), see Aliseda [2]. For a connectionist model of abduction, implemented in Bayesian networks, see Peng and Reggia [26].

[11]In adaptive logics, the type of abduction that has received most attention is indeed the simplest one: *abduction of a fact* [14].

of new connections. Discoveries of this sort represent the type of advance in science described by Kuhn as *Normal Science.*

On the other hand, with respect to the intuitive aspect of creativity, for Peirce scholars it is both well-known and controversial that in his view abduction is rooted in an animal instinct, which in turn involves the competence to conjecture hypotheses as a distinctive characteristic of the human capacity for truth approximation:

> In regard to instinctive considerations, I have already pointed out that 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. (CP [25]:7.220, my emphasis).

Thus the cognitive faculty that confers abduction on us is manifested as a rational instinct, as characterized by Ayim [8], or more precisely as the *guessing instinct*, the term used by Peirce himself. This hypothesis is similar to Chomsky's regarding the acquisition of language, according to which human ability for spoken language can be imputed to an innate competence. Nonetheless, Peirce's hypothesis has been sharply criticized, above all when it is used as a justification of the possibility of synthetic reasoning.[12]

Creativity combines as many intuitive as logical aspects. According to Peirce, the intuitive part is necessary for setting in train that innate capacity for guessing; but intuition is not the only aspect that counts for explaining success in the generation of abductive hypotheses. It is also necessary, although not by itself sufficient, that abductive hypotheses conform to a certain logical structure such as that provided by Peirce's formulation.

To conclude, I have argued in favor of the thesis that logic and creativity are not mutually exclusive, and have shown, through the abductive formulation of Peirce, that Logic has a privileged place in Creative Reason. I close with a citation from Simon, which captures very well the conception of creativity that I have chosen to defend in this text:

> New representations, like new problems, do not spring from the brow of Zeus, but emerge by gradual—and very slow—stages [...]. Most scientific activity goes on within the framework of established paradigms. Even in revolutionary science, which creates those paradigms, the problems and representations are rooted in the past; they are not created out of whole cloth. ([35]:301-302)

References

1. Aliseda, A., Gillies, D.: Logical, historical and computational approaches. In: Kuipers, T. (ed.) General Philosophy of Science: Focal Issues (Handbook of the Philosophy of Science), vol. 1, pp. 431–513. Elsevier, Amsterdam (2007)

[12]There are those who question it (Rescher [31]) and those who uphold it [6], but all agree in that the guessing instinct hypothesis presents serious problems.

2. Aliseda, A.: The logic of abduction: an introduction. In: Magnani, L., Bertolotti, T. (eds.) Handbook of Model-Based Science (Section The Logic of Hypothetical Reasoning, Abduction and Models). Springer, Heidelberg (2017a)
3. Aliseda, A.: El Lugar de la Lógica en la Razón Creativa. In: Pérez Ransanz, A.R., Ponce Miotti, A.L. (eds.) Creatividad e innovación en ciencia y tecnología (Colección del Seminario de Investigación sobre Sociedad del Conocimiento y Diversidad Cultural, 10), pp. 171–184. UNAM, México (2017b)
4. Aliseda, A.: Lógica como Herramienta de la Razón. Razonamiento Ampliativo en la Creatividad, la Cognición y la Inferencia. Cuadernos de lógica, epistemología y lenguaje, vol. 6. College Publications, Milton Keynes, U.K. (2014)
5. Aliseda, A.: Abductive Reasoning: Logical Investigations into Discovery and Explanation. Synthese Library, vol. 330. Springer, Dordrecht. (2006)
6. Anderson, D.: The evolution of Peirce's concept of abduction. Trans. Charles S. Peirce Soc. **22**(2), 145–164 (1986)
7. Anderson, D.: Creativity and the Philosophy of C.S. Peirce. Springer, Dordrecht (originally published by Martinus Nijhoff Publishers) (1987)
8. Ayim, M.: Retroduction: The rational instinct. Trans. Charles S. Peirce Soc. **10**(1), 34–43 (1974)
9. Campos, D.: On the distinction between Peirce's abduction and Lipton's inference to the best explanation. Synthese **180**, 419–442 (2011)
10. Douven, I. (2011). Abduction. In The Stanford Encyclopedia of Philosophy. Available via http://plato.stanford.edu/archives/spr2011/entries/abduction/. Accessed 5 Sept 2018
11. Echeverría, J.: Filosofía de la Ciencia. Akal, Madrid (1995)
12. Fann, K.T.: Peirce's Theory of Abduction. Martinus Nijhoff, The Hague (1970)
13. Frankfurt, H.: Peirce's notion of abduction. J. Philos. **55**(14), 593–597 (1958)
14. Gauderis, T.: Modelling hypothetical reasoning by formal logics. In: Magnani, L., Bertolotti, T. (eds.) Handbook of Model-Based Science, (Section "The Logic of Hypothetical Reasoning, Abduction and Models). Springer, Heidelberg (2017)
15. Hanson, N.R.: Observation and Explanation: A Guide to Philosophy of Science. Allen and Unwin, London (1971)
16. Harman, G.H.: The inference to the best explanation. Philos. Rev. **4**(1), 88–95 (1965)
17. Hookway, C.: Peirce. Routledge, London (1992)
18. Kapitan, T.: In What Way is Abductive Inference Creative? Transactions of the Charles S. Peirce Society. **26**(4), 499–512 (1990)
19. Kuipers, T.: From Instrumentalism to Constructive Realism: On Some Relations Between Confirmation, Empirical Progress, and Truth Approximation. Synthese Library 287. Kluwer Academic Publishers, Dordrecht (2000)
20. Lakatos, I.: Proofs and Refutations. Cambridge University Press, Cambridge (1976)
21. Langley, P., Simon, H., Bradshaw, G., Zytkow, J.: Scientific Discovery. Computational Explorations of the Creative Processes. MIT Press/Bradford Books, Cambridge (1987)
22. Laudan, L.: Why was the logic of discovery abandoned? In: Nickles, T. (ed.) Scientific Discovery, Logic and Rationality. Boston Studies in the Philosophy and History of Science, vol. 56, pp. 173–183. Kluwer Academic Publishers Reidel, Dordrecht (1980)
23. Musgrave, A.: Deductive heuristics. In: Gavroglu, K. et al. (eds.) Imre Lakatos and Theories of Scientific Change, pp. 15–32. Kluwer Academic Publishers, Dordrecht (1989)
24. Paavola, S.: Peircean abduction: instinct or inference?. Semiotica **153–154**, 363–374 (2005). Mouton de Gruyter, Berlin
25. Peirce, C.S.: Collected Papers (CP), vols. 1–8, Hartshorne, C., Weiss, P., Burks, A.W. (eds.). Harvard University Press, Cambridge, MA (1931–1958) (References to this work are written as (CP passage), following the international convention and observing Springer's format. The work compiles and orders Charles Peirce's writings in eight volumes.)
26. Peng, Y., Reggia, J.A.: Abductive Inference Models for Diagnostic Problem-Solving. Springer, New York (1990)

27. Popper, K.R.: The Logic of Scientific Discovery. Hutchinson, London/New York (first published in German 1935) (1959)
28. Pérez Ransanz, A.R.: Qué queda de la distinción entre contexto de descubrimiento y contexto de justificación? Theoria, Int. J. Theor. History Found. Sci. **22**(60), 347–350 (2007)
29. Reichenbach, H.: Experience and Prediction. University of Chicago Press, Chicago (1938)
30. Reichenbach, H.: The Rise of Scientific Philosophy. University of California Press, Berkeley (1951)
31. Rescher, N.: Peirce's Philosophy of Science. Critical Studies in His Theory of Induction and Scientific Method. University of Notre Dame Press, Notre Dame/London (1978)
32. Savary, C.: Discovery and its logic: Popper and the 'Friends of discovery'. Philos. Soc. Sci. **25**(3), 318–344 (1995)
33. Sharp, R.: Induction, abduction, and the evolution of science. Trans. Charles S. Peirce Soc. **6**(1), 17–33 (1970)
34. Simon, H.A.: Does scientific discovery have a logic? 326–337 (1977) (originally published in Philosophy of Science **40**, 471–480 (1973))
35. Simon, H.A.: Models of Discovery and Other Topics in the Methods of Science. Boston Studies in the Philosophy and History of Science vol. 54. Reidel, Dordrecht/Boston (1977)
36. Sternberg, R.J., Lubart, T.I.: The Concept of Creativity: Prospects and Paradigms. In: Sternberg, R.J. (ed.) Handbook of Creativity. Cambridge University Press (3rd reprint) (2003)
37. Thagard, P.R.: Computational Philosophy of Science. MIT Press, Cambridge (Bradford Books) (1988)
38. Thagard, P.R.: Conceptual Revolutions. Princeton University Press, Princeton (1992)
39. Thagard, P.R.: The unity of Peirce's theory of hypothesis. Trans. Charles S. Peirce Soc. **13**(2), 112–123 (1977)

On Abducing the Axioms of Mathematics

Woosuk Park

Abstract How do we discover and justify axioms of mathematics? In view of the long history of the axiomatic method, it is rather embarrassing that we are still lacking a standard answer to this simple question. Since the axiom of choice is arguably one of the most frequently discussed famous axioms throughout the history of mathematics, Thomas Forster's recent identification of the axiom as an inference to the best explanation (IBE) provides us with a nice point of departure. I will argue that, by separating sharply between abduction and IBE, we can give a convincing account of both the discovery and the justification of the axioms of mathematics.

1 Introduction: Forster on the Axiom of Choice and IBE

How do we discover and justify axioms of mathematics? In view of the long history of axiomatic method, it is rather embarrassing that we are still lacking a standard answer to this simple question.

Thomas Forster's recent paper "The Axiom of Choice and Inference to the Best Explanation" provides us with a nice point of departure:

> An argument often given for adopting the Axiom of Choice as an axiom is that it has a lot of obviously true consequences. This looks like a legitimate application of the practice of Inference to the Best Explanation. (Forster [6], 191)

Forster presents the argument as follows:

> There are various obviously true assertions, such as that the union of countably many countable sets is countable, or that a perfect binary tree has an infinite path, or that a finitely branching infinite tree must have an infinite branch, which — it transpires — cannot be proved without exploiting the axiom of choice. Given that they are obviously true, any sensible system of axioms for set theory will have to prove them, so we'd better include the axiom of choice in our set of axioms. (Forster [6], 191)

W. Park (✉)
Korea Advanced Institute of Science and Technology, Daejeon, South Korea
e-mail: woosukpark@kaist.ac.kr

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_8

And, he identifies the argument as a kind of abduction or inference to the best explanation (IBE):

> Although this argument sounds a bit like a fallacy of affirming the consequent, it's actually nothing of the kind. Arguments like this are so common and so natural and so legitimate that it is hardly surprising that this method has been identified by philosophers as a sensible way of proceeding and that there is a nomenclature for it and a literature to boot. It is probable that this is (at least part of) what Peirce had in mind when he coined the word abduction; nowadays it is captured by the expression inference to the best explanation "IBE"; see Lipton [27] for an excellent treatment. (Forster [6],192)

The axiom of choice is arguably one of the most frequently discussed famous axioms throughout the history of mathematics. The heated controversies around the discovery and justification of this axiom indeed supplies rich sources of insights for philosophers and historians of mathematics. Furthermore, by treating the crucial terms "abduction" and "inference to the best explanation (IBE)" as synonymous, Forster presents me a perfect target I intend to aim at in this paper. I will argue that, by separating sharply between abduction and IBE, we can give a convincing account of both the discovery and the justification of the axioms of mathematics.

Let me briefly outline what I intend to do in this paper. In Sect. 2, I shall discuss Russell's regressive method in mathematics, which has been extensively examined recently as a challenge to the received reading of Russell as one of the epistemic logicists. Whether it is just a historical accident or not, it is interesting to note that mathematicians and philosophers around David Hilbert were also extremely interested in the regressive method in mathematics. I will check wherein lie the similarities and the differences between their views and that of Russell's in Sect. 3. In Sect. 4, I will turn to problem of the discovery and the justification of the axiom of choice as a concrete example for comparing Russell's and the Zermelo's regressive method. In Sect. 5, I will discuss Gabbay and Woods' brilliant idea of distinguishing regressive abduction as abductive premise search and mathematical proof as enthymeme resolution. By emphasizing the distinction between abduction and IBE, however, I will speculate on the possibility of assimilating them by introducing the notion of abductive enthymeme in Sect. 6.

2 Russell's Regressive Method

According to the typical account, the emergence of foundations of mathematics was motivated by the crisis of mathematics due to Russell's paradox. This common epistemological concern can be easily understood in the case of Frege's logicism. For, as Irvine writes,

> Frege's original idea had been that if the principles of logic are granted self-evidence, and that if the laws of arithmetic can be shown to be derivable from them, arithmetic will have become epistemologically justified. (Irvine [20], 305)

One might think that Russell's logicism can be explained in basically the same fashion. Indeed, such a view baptized by Irvine as "epistemic logicism" has been the standard interpretation of Russell's logicism:

> According to this view, since we are epistemically justified in accepting the self-evident truths of logic, the successful identification of mathematics with logic would give us the same justification for accepting the truths of mathematics. (Ibid., p. 307)[1]

Irvine points out, however, epistemic logicism is not only inconsistent with "Russell's explicitly stated views on the subject" but also "susceptible to a number of well-known, related objections". (Ibid.)

The difference between Frege's epistemic logicism and Russell's own logicism may be hinted at by Russell's autobiographical remarks related to Russell's paradox (Russell [55], p. 58; Irvine [20], 314; Patton [43], p. 109). Irvine finds Russell's view "representative and telling", for it indicates that the discovery of Russell's paradox forced him "to develop the idea that much of logic would have to be justified by some method other than self-evidence". (Irvine [20], 314; Patton [43], p. 109)

What are the alleged well-known objections to epistemic logicism? Irvine aptly summarizes the basic idea of these objections as follows:

> It is unlikely that mathematics should gain its sole epistemic justification via logic since parts of mathematics are themselves more certain than (and are often known independently of) the requisite body of logical belief. (Ibid.)

Irvine counts objections based on this idea as "common sense objections", and presents the following as an example:

> For example, consider an elementary arithmetical proposition such as "$2 + 2 = 4$". What is it that provides the requisite justification for our acceptance of such a proposition? If it is suggested that our knowledge that such a proposition obtains results from a formal proof from first principles, immediate objections arise. Not everyone who knows such elementary propositions to be true, if pressed, could come up with a satisfactory sequence of sentences which would constitute the required proof. In fact, most people who know many elementary arithmetical truths could not even truthfully claim ever to have seen such a proof, let alone to being able to reproduce one as a warrant for their belief. [Ibid., 308]

According to Irvine, this kind of objections to logicism was actually raised by Poincaré [49, pp. 3f.] and Wittgenstein [61, §65^{e}ff.]. He also draws our attention to Steiner's detailed discussion of similar points (Steiner [59], p. 19; Irvine [20], 309).

It is interesting to note that Russell is well aware of these objections. Further, Irvine's claim that "Russell himself accepted many of them as sound" is fully supported by Russell's own words:

> There is an apparent absurdity in proceeding, as one does in the logical theory of arithmetic, through many rather recondite propositions of symbolic logic, to the "proof" of such truisms as $2 + 2 = 4$: for it is plain that the conclusion is more certain than the premises, and the supposed proof therefore seems futile. But of course what we are really proving is not the

[1] Irvine cites Steiner [59], Parsons [42], and Lakatos [25] as examples of such an interpretation. Also, he points out that this interpretation is indeed correct for other logicists such as Frege [7, p. 2f] and Hempel [13].

> truth of 2 + 2 = 4, but the fact that from our premises this truth can be deduced. (Russell [54; Irvine [20], 309)

As a consequence, many scholars have discussed recently "Russell's method of analysis and his regressive method for justifying the axioms of logic" as distinctively departing from Frege's logicism.[2] (Patton [43], p. 108). And, in all this discussion, Russell's talk "The Regressive Method of Discovering the Principles of Mathematics" to the Cambridge Mathematical Club on March 9, 1907 turns out to be the central text (Russell [54]).

Russell's aim in this talk is clearly announced by himself:

> My object in this paper is to explain in what sense a comparatively obscure and difficult proposition may be said to be a premise for a comparatively obvious proposition, to consider how premises in this sense may be discovered, and to emphasise the close analogy between the methods of pure mathematics and the methods of the sciences of observation. (Russell [54], 272; Irvine [20], 309, 323; Patton [43], p. 113)

Russell's distinction between two different types of premise, i.e., empirical premise and logical premise is introduced just for the first task, i.e., to explain in what sense a comparatively obscure and difficult proposition may be said to be a premise for a comparatively obvious proposition. Further, there is no doubt that Russell, only by fully employing this distinction, was not only able to present his original ideas related to the second task, i.e., how premises in this sense may be discovered, but also to pursue the third task, i.e., to emphasize the close analogy between the methods of pure mathematics and the methods of empirical sciences.

Let us see how Russell distinguishes between empirical and logical premises. He writes:

> Thus we see that the word 'premise' has two quite different senses: there is what we may call the 'empirical premise', which is the proposition or propositions from which we actually are led to believe the proposition in question; and there is what we will call the 'logical premise', which is some logically simpler proposition or propositions from which, by a valid deduction, the proposition in question can be obtained. Thus the empirical premises for 2 + 2 = 4 will be '2 sheep + 2 sheep = 4 sheep, and other like facts; whereas the logical premises will be certain principles of symbolic logic. (Russell [54], pp. 272–273)

The main purpose of Irvine's strenuous effort to clarify carefully what Russell intends by "empirical premise" and "logical premise" seems to be this. He wants to claim that "[m]ost often in mathematics the empirical and logical premises coincide". He believes, "[i]t is in exactly these cases that a mathematical proof is of direct epistemological value". [Ibid.] In other words, Irvine is here trying make sense of Russell's question as to "in what sense a comparatively obscure and difficult proposition may be said to be a premise for a comparatively obvious proposition?" by identifying "the mid-range concepts (concepts which are neither extremely fundamental nor extremely complex) that are commonly comprehended most readily". [Ibid.]

[2] Patton cites Irvine [21], Godwyn and Irvine [10], Hager [12], and Gandon [9] as examples.

Patton seems to be particularly helpful in understanding what is at stake here. For, as she claims, Russell's position is "that in the case of at least some mathematical axioms, our commitment to them as axioms need not require us to believe that they are true" (Patton [43], p. 108). But it may not be certain that Russell indeed took this position. Patton tries to appease some such worries by the following elaboration:

> This may seem un-logicist: should not the axioms be derivable from logical truths? But there is another way to think of the logicist method. It is possible to be a logicist, to consider that arithmetic is derivable from the Peano axioms, and *nonetheless to argue that some mathematical axioms need not be universal, univocal* a priori *truths, but can have a regressive justification derived from the method of analysis.* To a Russellian logicist, the results of mathematics can constitute evidence of a deeper set of logical relationships, and the clarification of these logical relationships can extend our knowledge, in mathematics and in logic. (Patton [43], pp. 108–109; Emphasis is mine.)

What exactly is the relation between the regressive method and the method of analysis? Patton, in this regard, seems to have in mind on Michael Beaney's careful discussion of the different modes of "analysis", as she quotes him as follows:

> we may distinguish three core modes of analysis… the regressive mode, concerned to identify the "starting-points" (principles, premisses, causes, etc.) by means of which something can be "explained" or "generated", the resolutive mode, concerned to identify the elements of something and the way they interrelate, and the interpretive mode, concerned to "translate" something into a particular framework (154). (Beaney [1], 154; Patton [43], p. 120)

Then, she claims that "Russell was committed to something one could call immanent logicism" by putting Irvine's and Beaney's results in broader context. What she means by "immanent logicism" is "a commitment to the view that analysis of logical structures within mathematics can lead to an expansion of knowledge". (Patton [43], p. 110)

What is important in Beaney's discussion of the different modes of analysis for Patton is that "[f]or Russell, distinguishing axioms from theorems is not just a matter of tracing back lines of dependence via the regressive method". (Patton [43], p. 121). Patton rightly counts the resolutive mode as "what is usually identified with logical analysis". We are rather familiar with this resolutive sense of "analysis", which seems to be working in our understanding of analysis in analytic philosophy. The point Patton wants to highlight is that, as Beaney observes, "both the regressive and resolutive modes presuppose the third:

> Any analysis presupposes a particular framework of interpretation, and work is done in interpreting what we are seeking to analyse – the analysandum – as part of the process of disclosing what is more fundamental. (Beaney [1], 155–156; Patton [43], *ibid*.)

Patton concurs with Beaney, for "only under a certain interpretation is one proposition more fundamental than another". Indeed, Patton emphatically claims that "Russell's regressive method for discovering the axioms of mathematics should be read in the larger context of the method of analysis" and of what she called "immanent logicism". As she claims, Russellian method of analysis is not just "the regressive discovery of the axioms on which certain inferences or results depend" but includes the analysis of "the logical structures of dependence of a theory, or of the fundamental logical conditions for mathematical results" (*Ibid*.).

3 Regressive Method in the Hilbertian Tradition

Patton's strategy to find the sources for Russell's regressive method by starting from his interlocutors, including Frege, Bradley, Keynes, and Poincaré turns out to be very successful. For, as she indicates, all these influential scholars not only discuss Jevons' *Principles of Science* [23] seriously, but also adopt many ideas from it. So, her historical conjecture that "the application of the "regressive method" to the axioms of mathematics and of empirical science was broadly established by the end of the nineteenth century" may indeed be true (Patton [43], pp. 110–113; See also Laudan [26]). By simply expanding the list, we may find further confirmations. For example, Peirce could have been mentioned by Patton (See Peirce [47, 48; See also articles in Moore [33]). It is more important, however, to note the striking fact that she seems to ignore Hilbert and his associates in her discussion of the history of the regressive method.

The overwhelming importance of Hilbert in the history of the axiomatic method cannot be too much emphasized. For, it is Hilbert who radically changed our conception of the axiomatic method itself. Further, we do have enough historical evidence that Hilbert and his associates at Göttingen were extremely interested in the regressive method in mathematics.

As Peckhaus points out, Hilbert was keenly aware of the necessity to distinguish sharply between "the presentation of systems of statements in axiomatic form" and "the axiomatic method", as is clear from his definition of the axiomatic method:

> I understand under the axiomatic exploration of a mathematical truth [or theorem] an investigation which does not aim at finding new or more general theorems being connected to this truth, but to determine the position of this theorem within the system of known truths in such a way that it can be clearly said which conditions are necessary and sufficient for giving a foundation of this truth. (Hilbert [14], 50; Peckhaus [46], 10)

After having characterized Hilbert's axiomatic method as "an architectural procedure" ordering conditions and propositions in mathemtics, Peckhaus claims that Hilbert later sharpened this idea of ordering in terms of "progressive and regressive tasks of mathematics:

> The progressive task consists in developing systems of relations and investigating their logical consequences. The regressive task consists in determining the conditions of a theory on the base of a clear distinction between suppositions and logical consequences. (Peckhaus [46], 10)

Hilbert's emphasis on the universality of the axiomatic method is salient in that it is not restricted to natural sciences but extended to other fields of knowledge such as economics or even to philosophy (Hilbert [16], p. 18; Peckhaus [46], 10–11). Indeed, Hilbert's pride of his unprecedented contribution to the axiomatic method can be highlighted by Hilbert's another catch phrase, i.e., "the deepening the foundations". In Hilbert's "Axiomatisches Denken" [15], where he emphasizes the central role of mathematics among the sciences, we read the following suggestive claims:

> The procedure of the axiomatic method, as it is expressed here, amounts to a deepening of the foundations of the individual domains of knowledge—a deepening that is necessary for every edifice that one wishes to expand and to build higher while preserving its stability. (Hilbert [15], p. 1109)

This sentence seems to claim that deepening the foundations of the individual science necessarily demands its expansion to its neighbouring fields of knowledge (Park [38, 39], p. 435).

Now we may note that many of Hilbert's associates seem to share Hilbert's views. To name a few evident examples, we may mention Zermelo, Bernays, and Nelson. Zermelo is the one who axiomatized set theory for the first time in 1908. Bernays is *the* collaborator of Hilbert, the true author of the monumental *Grundlagen der Mathematik, I and II* (Hilbert and Bernays [17]) co-authored with Hilbert. Nelson may not be widely known, though he has a unique place in the history of ethics. For this reason, Peckhaus' brief comments on Nelson could be useful:

> For Hilbert, the axiomatic method is a representation of mathematical thought, but again, its application is not restricted to mathematics. It can even be applied to sets of statements that cannot be based on axioms. Hilbert himself mentions as an example, the work of the Göttingen philosopher Leonard Nelson. Like Kant, Nelson strictly rejected attempts to apply the mathematical method, i.e., the axiomatic-deductive method, to philosophical questions. In his *Kritik der praktischen Vernunft* of 1917, however, he proposed to erect ethics by means of the axiomatic method. (Peckhaus [46], 11–12; See also Nelson [37])

From Reid's popular biography of Hilbert, we are rather well informed of Hilbert's effort to secure Nelson a stable academic position (Reid [52], pp. 144–145). Probably, however, what is important for my purpose is that Nelson was a leader of a group inspired by J. F. Fries. In fact, Zermelo and Bernays were also the members of this group (Reid [52], p. 151). As a consequence, we can gather the following information about Nelson from a biography of Zermelo:

> In cooperation with Gerhard Hessenberg, Nelson worked on a programme of "Critical Mathematics" ("Kritische Mathematik") as a philosophical alternative to Frege's logicism and Poincaré's anti-logicism. Already in 1904, when Nelson was still a student, both had become co-founders of the new series of *Abhandlungen der Friesschen Schule* which was to promote Kant's critical philosophy in the continuation of the philosopher Jakob Friedrich Fries (1778–1843) (Ebbinghaus and Peckhaus [5], p. 72; Cf., e.g., Nelson [36]). For details see Peckhaus [44], p. 123f.

Now we can see that Nelson advocates critical mathematics in the sense of Fries, who admired Kant so much. According to Pulte's recent assessment, Fries tries to develop Kantian philosophy further by revision and extension. It is simply beyond the scope of this paper to discuss in what ways Fries tried to go beyond Kant. For my present purpose, what is important is that he made pivotal influence to the New Friesian School, and thereby to the Hilbetians. [For further details, see Peckhaus [44, 45] and Pulte [50], p. 46] Further, as Peckhaus points out, Hilbert's regressive method itself might have its origin in Kantian notion of "regressive method" (Peckhaus [46] 11–12; See Kant [24], A 42, § 5, note).

4 How Can the Axiom of Choice Be Smuggled?[3]

4.1 Russell

Patton starts her discussion of Russell's involvement with the axiom of choice by a seemingly neutral description:

> In the early 1900s, Russell became increasingly aware of the role of certain axioms, including the axioms of reducibility, choice, and infinity, and the multiplicative axiom, in deriving the results of mathematics. (Patton [43], p. 116)

From her point of view, the axiom of choice is just one of the axioms of mathematics, insofar as the issue is the discovery and justification of the axioms of mathematics. Further, she is emphasizing the role of these axioms "in deriving the results of mathematics" rather than establishing their truth. In other words, she already characterizes Russell's position in terms of his regressive method, thereby distancing it from epistemic logicism. Probably, for this reason, she is able to read Moore's general remarks on the role of the Axiom of Choice in Russell's work as supporting her perceptive reading of the title of Russell's talk, "The Regressive Method of Discovering the Principles of Mathematics". Moore simply writes:

> The role played by the Axiom of Choice in Russell's work is long and complicated. Like many mathematicians around 1900, he used the Axiom of Choice implicitly without realizing that he had done so, since no one had yet formulated it explicitly (lv). (Moore [34], p. Lv.)

But Patton counts him as putting in his own way her insight that, unlike epistemic logicists, Russell's regressive method is "to identify the *principles* of mathematics that were always at work in, and implicit in, mathematical practice" rather than "to discover *the* principles of mathematics: the single, fundamental, true laws of logic from which mathematics is derived" (Patton [43], p. 117; Patton's emphasis).

In order to explain how Russell was led to his own and Whitehead's Multiplicative axiom, which is logically equivalent to the Axiom of Choice, Patton again relies on Moore's exposition.[4] She quotes Moore's description of Russell's route to the Multiplicative Axiom as follows:

> Russell reached the Multiplicative Axiom by a route quite different from Zermelo's path to the Axiom of Choice. While Zermelo had been preoccupied with the Well-Ordering Problem, Russell's Multiplicative Axiom arose from considering the infinite product of disjoint sets, i.e., their multiplicative class, in order to define the product of infinitely many cardinals. (Moore [35], 123; Patton [43], pp. 117–118)

Based on Moore's nice report, Patton also emphasizes the following fact:

---

[3]I heavily rely on Moore [35] as the basic data for my discussion, as Maddy and Patton do (Maddy [29, 30]; Patton [43]). See also Bell [2] and Jech [22].

[4]Even in quoting from Russell's letter to Jourdain (Letter of March 15, [53]; reprinted in Grattan-Guiness [11], 107), she is again following Moore's lead (Patton [43], p. 117; Moore [35], p. 122).

> As Moore notes, at first, Russell and Whitehead saw the Multiplicative Axiom as a theorem. At some point, Russell came to realize that a number of results depended on the proposition and came to regard it as an axiom. (Patton [43], p. 118)

But, exactly in what respects Russell's route to his and Whitehead's Multiplicative Axiom is different from Zermelo's route to the axiom of Choice?

4.2 Zermelo

Zermelo's route to the Axiom of Choice was via the Well-Ordering Problem: Can every set be well-ordered? Here again let us follow Moore's guide:

> While the Axiom's roots lie in the infinite processes so prominent in nineteenth-century analysis, this assumption was not explicitly stated until a new discipline emerged from analysis: Cantorian set theory. The decisive question, here termed the Well-Ordering Problem, was this: Can every set be well-ordered? Strictly speaking, this Well-Ordering Problem did not originate as a problem at all, but as an assumption. In 1883, Georg Cantor proposed as a valid "law of thought" that every set can be well-ordered. (Moore [35], p. 4)[5]

As Moore continues to explain,

> Yet this law of thought, hereafter termed the Well-Ordering Principle, was not accepted by his contemporaries, and a decade later Cantor himself was seeking a proof. It was in order to obtain such a proof that in 1904 Ernst Zermelo first formulated the Axiom of Choice. (Moore [35], p. 2)

Now, let us probe the question as to what exactly is different in Russell's case and Cantor's and Zermelo's case independently of how other mathematicians and philosophers found the matter. Interestingly enough, there seem to be some obvious differences between them. Russell thought that the Multiplicative Axiom is a theorem. Only after having failed to prove it, he began to count it as an axiom. On the other hand, Cantor and Zermelo never thought the Well-Ordering Principle as a theorem. Cantor thought it to be an assumption. But later he tried to prove it. In other words, he became to count it as a theorem. And, Zermelo formulated the Axiom of Choice to prove it. That means, Zermelo also treated the Well-Ordering Principle as a theorem to be proven.

Moore's report is continued as follows:

> In this manner, the Axiom passed from unconscious to conscious use and, for many mathematicians at the time, to conscious avoidance. Indeed, Zermelo's solution to the Well-Ordering Problem generated a heated debate, which raged among mathematicians all over Europe, as to whether his Axiom and proof were correct. This controversy soon led him to publish, on the one hand, a spirited defense of his proof and, on the other, his axiomatization for set theory. Within this axiomatization, he believed, both his Axiom and proof were securely embedded. His critics did not agree. (Moore [35], p. 2)

[5] Please refer to my more extensive discussion of Zermelo in Chap. 5 of Park [40].

Moore's description of what happened in mathematical community in response to Zermelo's proof of the Well-Ordering Principle and his axiomatization of set theory is too short to be informative. Yes, there was a heated controversy, and his critics never agreed with Zermelo's belief that both the Axiom of Choice and the proof of the Well-Ordering Principle were securely embedded within his axiomatization of set theory. But exactly why should there be such a controversy? For what reason didn't the critics agree with Zermelo's view?

Moore himself must have struggled to understand what happened. And, he seems to have arrived at the following conclusions.

> M1: The use of infinitely many arbitrary choices had been growing independently during the last third of the nineteenth century.
>
> M2: But there was no awareness that a new mathematical principle was required.
>
> M3: It was only after Zermelo's proof of the Well-ordering Principle within his axiomatic set theory that mathematicians became sensitive to the theorems that need the Axiom of Choice to be proven. (Moore [35], p. 2)

At first blush, it is apparently a serious lacuna that Cantor is not mentioned at all in any of M1–M3. However, Moore claims that "Cantor did more than this" (Moore [35], p. 30–31; See also Zermelo in Cantor [4], p. 451). Moore seems to be claiming here:

> M4: Cantor unconsciously employed the Assumption (i.e., the Axiom of Choice) in certain basic results (rather than in set theory and point-set topology).
>
> M5: Only through its own offsprings, the Assumption came to be used by mathematicians. (Moore [35], pp. 30–31, p. 38)

5 Woods on Regressive Abduction and Enthymeme Resolution

In order to understand better the difference between Zermelo and Russell in handling the Axiom of Choice, I suggest to employ Dov Gabbay and John Woods' comparison of the elements of abduction in the regressive premise search and in the enthymeme resolution. Even though the abductive element in Russell's regressive method has been detected by some philosophers (Irvine [20], 322, n. 26; Patton [43], p. 113), only Gabbay and Woods use the expression "regressive abduction" (Gabbay and Woods [8], p. 125f.). Gabby and Woods are also interested in "the problem of determining in what sense 'a comparatively obscure and difficult proposition may be said to be a premise for a comparatively obvious proposition' (Russell [54], p. 272), and the problem of "explaining how such obscure and difficult propositions are discovered and justified" (Gabbay and Woods [8], p. 125).

Gabbay and Woods schematize regressive abduction as follows:

> Let T be the task of justifying a non-obvious ("recondite" is Russell's word for it) principle of logic. Call this proposition H. Suppose that there is an obvious truth of mathemtics, V, such that for some set of propositions K, $K(H)$ constitutes a proof of V. Since nothing else

(so far as one can see) counts in favour of the truth of H, H is forwarded conjecturally soley on the basis of the role it plays in the proof of Q. (Gabbay and Woods [8], p. 297)

Against this background, Gabbay and Woods contrast Wiles' situation in proving Fermat's Last Theorem. It is a brilliant idea to characterize what Wiles had to do as solving Fermat's enthymeme:

Then Wiles' T' was to construct a proof of V'. His agenda was both similar to and different from that undertaken by regressive abducers. What the regressive abducer also wants is a proof of V. But that is not all that he wants, or even the most important part of it. What he wants is a proof of V that contains H as a prior line. The regressive abducer's main task is to justify H, not prove V. Proving V is a subagenda whose whole motivation is the support it lends not to V but to H. Wiles, on the other hand, was wholly absorbed with V'. To that end, there are thousands of prior lines, towards none of which did Wiles' show the kind of interest that defines regressive abduction. Wiles took great pains to construct his derivation of V' from mathematically sound premisses (and subproofs), but nowhere in that vast undertaking is there the slightest indication that Wiles' wanted his proof of Fermat's theorem to constitute a justification of any of the premisses he actually deployed. Regressive abducers target H_s. Enthymeme resolvers target V'. (Gabbay and Woods [8], pp. 297–298)

Let us distinguish between Russell$_t$ (Russell during the period in which he believed the Multiplicative Axiom to be a theorem) and Russell$_a$ (Russell during the period in which he believed it to be an assumption). Russell$_t$ was an enthymeme resolver targeting V'. On the other hand, it may not be easy to describe what Russell$_a$ was doing. Why? Because it is clear how Russell would discover and justify the Multiplicative Axiom. Likewise, we may distinguish between Cantor$_a$ (Cantor during the time in which he believed the Well-Ordering principle to be an assumption) and Cantor$_t$ (Cantor during the period in which he believed it to be something to be proven). Can we say without hesitation that Cantor$_a$ was a regressive abducer? Can we present unreservedly Cantor$_t$ as an enthymeme resolver? I think that both are rather difficult problems. Cantor$_a$ thought that the Well-ordering Principle must be self-evidently true. So, it is difficult to count him as a regressive abducer. Since he was interested in establishing theorems of set theory, he must be an enthymeme resolver insofar as those theorems are concerned, but not for the Well-Ordering Principle or the Axiom of Choice.

Probably, Zermelo may be the toughest case to handle with Gabbay and Woods' frmework. Zermelo$_t$ (Zermelo in the period in which he tried to prove the Well-Ordering Principle) may be an enthymeme resolver. However, he must have become a regressive abducer when he began to build an axiomatic system for set theory. Was there any one moment or a period of time at which he was both a regressive abducer and an enthymeme resolver at the same time in connection with one of his axioms of his set theory? How about regarding the Axiom of Choice?

6 Concluding Remarks

As was hinted at in the previous section, Gabbay and Woods' distinction between an enthymeme resolver and a regressive abducer seems to have an enormous potential to enhance our understanding of the problem of the discovery and the justification of axioms of mathematics. In liew of a conclusion, I would like to speculate a bit more about the ways of expanding their distinction. What would happen if we introduce the idea of enlarging the notion of enthymeme to cover not only the deductive enthymemes but also non-deductive enthymemes? Gabbay and Woods' observation that enthymeme resolution does not have the character of regressive abduction may be correct. But if we expand the scope of enthmeme to cover non-deductive as well as deductive enthymemes, then some cases of enthymeme resolution may or may not have the character of regressive abduction.

Wiles was an enthymeme resolver rather than a regressive abducer in trying to prove Fermat's Last Theorem. Thanks to Kenneth A. Ribet's proof inspired by Gerhard Frey, it turns out that proving Taniyama-Shimura Conjecture would prove Fermat's Last Theorem at the same time. [See Wiles [60], 443; Singh [57], p. 202; Singh and Ribet [58], 71–72]. As Gabbay and Woods claim, Wiles was exclusively interested in proving the truth of Taniyama-Simura Conjecture, and thereby that of Fermat's Last Theorem without bothering to find any fundamental propositions from which Taniyama-Simura Conjecture can be deduced. But couldn't we be able to view Frey (or Ribet) as both an enthymeme resolver and regressive abducer?

At the beginning of this paper, we saw Forster counting abduction and IBE as synomyms. However, Gerhard Minnameier attempts to contrast abduction with IBE in terms of their functions: "Thus, while abduction marks the process of generating theories—or, more generally, concepts—IBE concerns their evaluation" [32]–76]. Minnameier's opinion is widely shared, e.g., by Magnanai [31], Campos [3], and Mackonis [28]. [Park [39], pp. 29–31] Also, I pointed out that IBE cannot be abduction in GW-Model. [ibid., 31–34].[6]

Armed with both Gabbay and Woods' distinction between an enthymeme resolver and a regressive abducer and the widely held distinction between abduction and IBE, now we may compare what exactly Frey, Ribet and Wiles were doing in connection with the Fermat's Last Theorem. At a symposium in Oberwolfach in the autumn of 1984, Gerhard Frey made a remarkable claim that "if mathematicians could prove the Taniyama-Shimura conjecture then they would automatically prove Fermat's Last Theorem". [Singh [57], pp. 194–195]

The significance of Frey's truly creative insight to connect Taniyama-Shimura conjecture with Fermat's Last Theorem cannot be too much emphasized:

[6] I also discussed this problem in Park [41]. I am delighted by Raftopoulos's reminder that Hintikka [18, 19] already argued for the distinction between abduction and IBE (Raftopoulos [51], 262–263). In fact, Woods also discuss the problems related to abduction and IBE in his writings. For example, he discusses them in connection with the problem of abductive premiss-searches (Woods [62], p. 2, [63], p. 136). Woods' most recent attempt to prove the impossibility of identifying abduction with IBE is discussed in Park [41].

> For the first time in a hundred years the world's hardest mathematical problem looked vulnerable. According to Frey, proving the Taniyama-Shimura conjecture was the only hurdle to proving Fermat's Last Theorem. [Ibid., p. 197]

As is well known by now, there was "an elementary blunder" in Frey's logic: "he [Frey] had not quite demonstrated that his elliptic equation was sufficiently weird". [Ibid.] It was Ken Ribet who "proved that the Taniyama-Shimura conjecture implies Fermat's Last Theorem" in 1986. [ibid., p. 201] Only on hearing the news about Ribet's proof, Wiles plunged into his serious adventure to prove the Taniyama-Shimura conjecture, thereby to prove Fermat's Last Theorem ultimately. [Ibid., p. 205]

For my present purpose, i.e., to compare Frey, Ribet, and Wiles as enthymeme resolvers and/or regressive abducers, ample information is already at hand. As Gabbay and Woods seem to believe, Wiles might be characterized as merely an enthymeme resolver. For he was at least primarily interested in proving the Taniyama-Shimura conjecture. Frey and Ribet were, unlike Wiles, not only enthymeme resolvers but also regressive abducers, insofar as they bothered to demonstrate the relation between the Taniyama-Shimura conjecture and Fermat's Last Theorem. From my point of view, Frey is much more interesting than Ribet both as an enthymeme resolver and a regressive abducer. While Ribet was a successful enthymeme resolver in the usual sense of enthymeme, Frey can be called an enthymeme resolver only with some qualifications. As was pointed out above, he did not quite prove that the Taniyama-Shimura conjecture implies Fermat's Last Theorem. At best, he proved merely the plausibility of that implication. So, he was an unsuccessful enthymeme resolver: he neither turned the invalid deductive argument into a valid one, nor did he show the impossibility of doing so. As a regressive abducer, Frey and Ribet were not on a par. It was only Frey who detected the possibility of the Taniyama-Shimura conjecture's implying Fermat's Last Theorem. All too probably both Frey and Ribet might have tried any number of inferences to the best explanation in their attempts to prove that the Taniyama-Shimura conjecture implies Fermat's Last Theorem. So, Ribet was a regressive abducer only in the sense of doing an IBE or simply following the step of Frey's abduction.

References

1. Beaney, M.: Russell and Frege. In: Griffin, N. (ed.) The Cambridge Companion to Russell. Cambridge: Cambridge University Press (2003)
2. Bell, J.L.: The Axiom of Choice. College Publications, London (2009)
3. Campos, D.G.: On the distinction between Peirce's abduction and Lipton's inference to the best explanation. Synthese **180**, 419–442 (2011)
4. Cantor, G.: Gesammelte Abhandlungen mathematischen und philosophischen Inhalts. In: E. Zermelo (ed.) Springer, Berlin (1932), reprinted Olms, Hildesheim (1962)
5. Ebbinghaus, H.-D., Peckhaus, V.: Ernst Zermelo: An Approach to His Life and Work. Springer, Berlin/Heidelberg (2007)
6. Forster, T.: The axiom of choice and inference to the best explanation. Logique et Analyse **194**, 194–197 (2006)

7. Frege, G.: The Foundations of Arithmetic, 2nd ed. trans. Austin. Blackwell, UK (1884/1974)
8. Gabbay, F., Woods, J.: The Reach of Abduction: Insight and Trial. Elsevier, Amsterdam (2005)
9. Gandon, S.: Russell's Unknown Logicism. Palgrave McMillan, London (2012)
10. Godwyn, M., Irvine, A.: Bertrand Russell's Logicism, pp. 171–201 in Griffin (2003)
11. Grattan-Guinness, I.: Bertrand Russell on his paradox and the multiplicative axiom. An unpublished letter to Philip Jourdain. J. Philos. Logic **1**(2), 103–110 (1972)
12. Hager, P.: Russell's method of analysis, pp. 310–331 in Griffin (2003)
13. Hempel, C.: On the nature of mathematical truth. Am. Math. Monthly **52**, 543–556 (1945)
14. Hilbert, D.: ¨Uber den Satz von der Gleichheit der Basiswinkel im gleichschenkligen Dreieck. Proc. London Math. Soc. **35**, 50–68 (1902/03)
15. Hilbert, D.: Axiomatic thought. In: ed. by Ewald, W. (ed.) From Kant to Hilbert: A Source Book in the Foundations of Mathematics, vol. 2, Oxford, Clarendon Press, 1996, pp 1105–1115 (1918)
16. Hilbert, D.: Natur und mathematisches Erkennen. In: Rowe, D.E. (ed.) Birkhäuser Verlag, Basel/Boston/Berlin (1992)
17. Hilbert, D., Bernays, P.: Grundlagen der Mathematik I, II. Springer, Berlin/Heidelberg (1934, 1939)
18. Hintikka, J.: What was Aristotle doing in his early logic, anyway?: A reply to Woods and Hansen. Synthese **113**, 241–249 (1997)
19. Hintikka, J.: "What is abduction? the fundamental problem of contemporary epistemology", Trans. Charles S. Peirce Soc. **34**(3), 503–533 (1998)
20. Irvine, A.: Epistemic logicism & Russell's regressive method. Philos. Stud. **55**(3), 303–327 (1989)
21. Irvine, A.: Bertrand Russell's logic. In: Gabbay, Woods (eds.) Handbook of the History of Logic, vol. 5, pp. 1–28. Elsevier, Amsterdam (2009)
22. Jech, T.J.: The axiom of choice. Dover Publications INC, Mineola, N. Y. (1973)
23. Jevons, W.S.: Principles of Science. Richard Clay & Sons, London (1874)
24. Kant, I.: Prolegomena zu einer jeden künftigen Metaphysik dit als Wissenschaft und auftreten können. Johann Friedrich Hartknoch, Riga (1783)
25. Lakatos, I.: Infinite regress and foundations of mathematics. In: Worrall and Currie (eds.) Mathematics, Science and Epistemology, vol. 2, pp. 3–23. Cambridge University Press, Cambridge (1962/1980)
26. Laudan, L.: Theories of scientific method from Plato to mach. Hist. Sci. **7**, 1–63 (1968)
27. Lipton, P.: Inference to the Best Explanation, 2nd edn. Routledge, London (2004)
28. Mackonis, A.: Inference to the best explanation, coherence and other explanatory virtues. Synthese **190**(6), 975–995 (2013)
29. Maddy, P.: Realism in Mathematics. Oxford University Press, Oxford (1990)
30. Maddy, P.: Naturalism in Mathematics. Oxford University Press, Oxford (1997)
31. Magnani, L.: Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Ewasoning. Springer, Berlin (2009)
32. Minnameier, G.: Peirce-suit of truth—why inference to the best explanation and abduction ought not to be confused. Erkenntnis **60**, 75–105 (2004)
33. Moore, M.E.: New Essays on Peirce's Methematical Philosophy. Open Court, Chicago and La Salle, Illinois (2010)
34. Moore, G.H.: "Introduction" to Russell, Bertrand. In: Moore, G.H. (ed.) The Collected Papers of Bertrand Russell, Volume 5: Toward Principia Mathematica, pp. 1905–1908. Routledge, London (2015)
35. Moore, G.: Zermelo's Axiom of Choice: Its Origins, Development, and Influence. Springer, New York (1982, 2012)
36. Nelson, L.: Kritische Philosophie und mathemtische axiomatik. Unterrichtsblätter für Mathemtik und Naturwissenschften **34**, 108–142 (1928)
37. Nelson, L.: Kritik der praktischen Vernunft. Göttingen: Verlag "Öffentliches Leben" (1917)
38. Park, W.: Friedman on implicit definition. in search of the Hilbertian heritage in philosophy of science. Erkenntnis **76**(3), 427–442 (2012)

39. Park, W.: Abduction in Contect: The Conjectural Dynamics of Scientific Reasoning. Springer, Berlin (2016)
40. Park, W.: Philosophy's Loss of Logic to Mathematics. Springer, Berlin (2018)
41. Park, W.: Hintikka/Woods and Hansen Controversy on Aristotelian Fallacies. In: Gabbay, D., Magnani, L., Pietarinen, A., Park, W. (eds.) Natural Arguments: A Tribute to John Woods. College Publications, London, forthcoming (2019)
42. Parsons, C.: "Mathematics, Foundations of. In: Edwards, P. (ed.) The Encyclopedia of Philosophy, vol. 5, pp. 188–213 Macmillan and The Free Press, New York and London (1967)
43. Patton, L.: Russell's method of analysis and the axioms of mathematics. In: Lapointe, S., Pincock, C. (eds.) Innovations in the History of Analytical Philosophy. Palgrave macmillan, London (2017)
44. Peckhaus, V.: Hilbertprogramm und Kritische Philosophie. Das Göttinger Modell interdisziplinärer Zusammenarbeit zwischen Mathematik und Philosophie. Vandenhoeck & Ruprecht, Göttingen (1990)
45. Peckhaus, V.: Fries in Hilbert's Gottingen: Die Neue Fries'sche Schule. In: Hogrebe Wolfram and Herrmann Kay (eds.) Jakob Friedrich Fries—Philosoph, Naturwissenschaftler und Mathematiker, pp. 353–386. Lang, Frankfurt a.M (1999
46. Peckhaus, V.: Regressive analysis. In: Logical Analysis and History of Philosophy, 5 (2002). Available at http://www.uni-paderborn.de/fileadmin/kw/Institute/Philosophie/Personal/Peckhaus/Texte_zum_Download/regressive_analysis.pdf)
47. Peirce, C.S.: Collected Papers, 8 vols. In: Hartshorne, C., Weiss, P. (vols. I–VI), Burks, A.W. (vols. VII–VIII) (eds.) Harvard University Press, Cambridge, MA (1931–1958)
48. Peirce, C.S.: The New Elements of Mathematics, vol. 4. In Eisele, C. (ed.). Mouton de Gruyter, Berlin/New York; Atlantic Highlands, Humanities Press, NJ (1976)
49. Poincare, H.: Science nd Hypothesis (1905). Dover, New York (1952)
50. Pulte, H.: J. F. Fries' philosophy of science, the New Friesian School and the Berlin Group: on divergent scientific philosophies, difficult relations and missed opportunities. In: Milkov, N., Peckhaus, V. (eds.) The Berlin Group and the Philosophy of Logical Empiricism. Springer, Dordrecht (2013)
51. Raftopoulos, A.: Abduction, inference to the best explanation, and scientific practice: the case of Newton's optics. In: Magnani, L., Casadio, C. (eds.) Model-Based Reasoning in Science and Technology, pp. 259–277. Springer (2016)
52. Reid, C.: Hilbert. Springer, Berlin/New York (1970, 1996)
53. Russell, B.: M. Poincaré's Science et Hypothèse. Mind **15**(57), 141–143 (1906)
54. Russell, B.: The regressive method of discovering the premises of mathematics. In: Lackey, D. (ed.) Essays in Analysis, pp. 272–283. George Allen & Unwin Ltd, London (1907/1973)
55. Russell, B.: My Philosophical Development. Uniwin Books, London (1959/1975)
56. Russell, B.: The Collected Papers of Bertrand Russell, Volume 3. In: Moore, G. (ed.). Routledge, New York (1993)
57. Singh, S.: Fermat's Enigma. Anchor Books, New York (1997)
58. Singh, S., Ribet, K.A.: Fermat's last stand. Sci. Am. **277**(5), 68–73 (1997)
59. Steiner, M.: Mathematical Knowledge. Cornell University Press, Ithaca and London (1975)
60. Wiles, A.: Modular elliptic curves and Fermat's Last Theorem. Annals of Mathematics **142**, 443–551 (1995)
61. Wittgenstein, L.: Bemerkungen über die Grundlagen der Mathematik, trans. By G. E. M. Anscombe as Remarks on the Foundations of Mathematics, Anscombe, G.E.M., Rhees, R., Von Wright, G.H. (eds.). Blackwell, Oxford (1956)
62. Woods, J.: Aristotle's Earlier Logic. Hermes Science Publishing Ltd, Oxford (2001)
63. Woods, J.: Aristotle's Earlier Logic, 2nd edn. College Publictions, London (2014)

Abduction, Complex Inferences, and Emergent Heuristics of Scientific Inquiry

John R. Shook

Abstract The roles of abductive inference in dynamic heuristics allows scientific methodologies to test novel explanations for the world's ways. Deliberate reasoning often follows abductive patterns, as well as patterns dominated by deduction and induction, but complex mixtures of these three modes of inference are crucial for scientific explanation. All possible mixed inferences are formulated and categorized using a novel typology and nomenclature. Twenty five possible combinations among abduction, induction, and deduction are assembled and analyzed in order of complexity. There are five primary categories for sorting these inferential procedures: fallacies, non-scientific procedures, quasi-scientific procedures, scientific procedures, and scientific heuristics.

Experimental sciences use abductions in the course of their methodologies. The involvement of abductive inferences in many kinds of dynamic heuristics allows scientific methodologies to consider and test novel explanations for curious matters, and to gradually increase information about the world's ways.

Science didn't invent abductive inference; it was borrowed. Deliberate reasoning in general—accepting conclusions due to their discerned relationships with relied-upon beliefs—frequently follows abductive patterns as well as deductive and inductive patterns. Deeper cognitive processes such as perception, concept formation, and shifting habits of thought likely modes of abduction.[1] To the extent that the experimental sciences contribute increases in knowledge, they have applied some

J. R. Shook (✉)
University at Buffalo, Buffalo, NY 14260, USA
e-mail: jrshook@buffalo.edu

[1]Ordinary abduction is evidently habitual in practice for humans, and habits can be brought under reflective review for deliberation, especially if they are acquired in learning (Magnani [1]). No "instinct or inference" dichotomy about abduction is forced upon us, as if learning must be rigid and automatic. It is a debatable question whether something akin to abduction is instinctive for non-human animals (Park [2]).

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_9

abduction as well as induction and deduction. The power of scientific explanation does not reside within any of these inferential modes alone. Understanding roles for abduction in satisfactory explanation should look to complex mixtures of these three modes.

A preliminary exploration of those modes, organized by increasing complexity and categorized with a typology, maps out some prominent features of this inferential landscape. Five primary categories emerge for the twenty-five combinatorial possibilities among deduction, induction, and abduction. These inferential procedures can be sorted into fallacies, non-scientific procedures, quasi-scientific procedures, scientific procedures, and scientific heuristics. Among these procedures are core methodologies inherent to metaphysical and theological worldviews, and their accurate classification helps to reveal their close relationships with proto-scientific thinking.

1 Why Abduction?

Deductive reasoning alone may be sufficient to intelligibly relate all knowledge already possessed. Alan Musgrave defends deductivism by pointing out how any generalization appearing to arise from non-deductive reasoning can be re-cast afterwards as a deductive inference with just the right premises added. After showing how to do this with a typical form of generalization, he adds,

> "The same applies to all the other patterns of inductive or ampliative reasoning. All can be reconstructed as deductive arguments with suppressed factual or epistemic premises."[2]

Reconstructed deductive arguments are useful in their own way, after new information has been established. Of course, knowledge arises (for humans, at least) from learning, and we must figure out which factual or epistemic premises are just the right ones. Being told that some extra fact about the world will convert non-deductive support for a conclusion into its deductive support isn't helpful for learning. Learners want to acquire precisely which fact, when supposed, will turn out to be the right support.[3]

Inquiry crucially relies on abduction, so that proposed matters can become believable supposed facts. Abduction, by itself, is a blatant fallacy—yet there appears to be no way to avoid it. Neither deduction (necessary inference) nor induction (probable inference) can increase the real amount of information beyond what is already accepted, but abduction (possible inference) can.

If these are the three primary modes of inference, with abduction playing a needed but insufficient role, then abduction may be transcending mere fallacy through its application in concert with deduction and/or induction. Charles Peirce, abduction's

[2] Musgrave [3, p. 127].

[3] Stanford [4] makes a similar point regarding the supposed self-sufficiency of Bayesian confirmation.

'discoverer', typically situated abduction alongside deduction and abduction in the proper functioning of scientific inquiry. His 1903 Harvard Lectures on Pragmatism is an example:

> Abduction merely suggests that something may be. Its only justification is that from its suggestion deduction can draw a prediction which can be tested by induction, and that, if we are ever to learn anything or to understand phenomena at all, it must be by abduction that this is to be brought about.[4]

Peirce placed immense confidence in abduction's explanatory powers, so long as it played a helpful role in cooperation with the other modes of inference.[5]

Here, we explore how patterns of procedural abduction—combinations of these three inferential modes executed over time towards some conclusion—can simultaneously reduce the fallacious character of abductions, increase the credibility deserved by their conclusions, and yield increasing information about the world. This preliminary work is concerned with methodologies within empirical inquiry, not about the overall strengths and weaknesses to "inference to the best explanation" or debates over scientific realism.[6] It delineates, identifies, and evaluates many combinations of deduction, induction, and abduction, from simpler forms to quite complex patterns. Some intricate combinations rise to the level of utility for experimental scientific inquiry. Along the way, non-scientific and pseudo-scientific procedures are exposed as well, which suggests why they can be relevant to the perennial demarcation problem.

Procedural abduction plays a significant role within some phases of proto-scientific and fully scientific methods. Its patterns may be most recognizable in the ordinary inquiries people undertake daily to sort and select simple explanations behind encountered events. Understanding the merits and risks inherent to procedural abduction would not be out of place in an effort to improve critical thinking. Science is by no means 'common sense' enlarged, although continuities are present.

It must be firmly noted from the outset that these patterns cannot be the "essence" of scientific methodology, if there could be such a thing. Nor are these patterns even capable of characterizing the more important inferential methods applied in the sciences. Some of the most complex abductive patterns do begin to resemble what have been called 'heuristics' to scientific inquiry, as later sections note. Science surely gets vastly more complicated than even the most convoluted inferential patterns categorized here. Nevertheless, in these procedural abductions, some of science's proto-methodical 'building blocks' can be discerned, and distinguished from inferential patterns on paths tending to diverge away or run parallel to science.

---

[4]Peirce [5, para. 171–172].

[5]A handful of recent philosophers of science have appreciated Peirce and abduction's significant role. Consult for example McMullin [6], Niiniluoto [7], Psillos [8]. A recent examination of Peirce's mature logic of scientific methodology is by Pietarinen and Bellucci [9].

[6]Campos [10] distinguishes Peirce's abduction apart from inference to the best explanation. For broader explorations of abduction's role in procedures of explanatory reasoning, consult Flach and Kakas [11], Lipton [12], Paavola [13], Aliseda [14], Pizzi [15], Schurz [16], Gauderis and De Putte [17], Gauderis [18], Aliseda and Beirlaen [19], Velázquez-Quesada [20].

2 Abduction Basics

Begin with abduction in its simplest form:

Q

If A then Q

So, A [“Simple abduction” – the ‘affirming the consequent’ fallacy]

Because we are only considering deliberate inferences, and such cognitions are extended in time, this three-part inference is to be understood as displaying temporal phases that matter to the acceptability of any conclusion. Here, ‘Q’ is learned first, and ‘If A then Q’ is considered after Q is already in mind, with the conclusion following in both temporality and plausibility (if any may pertain). After acquaintance with all three parts, they may be kept in mind as a single whole for further consideration, just as all the notes of a simple tune may be sustained together in the imagination without having to sing it over and over. Just as the original order of a tune’s notes still matter all the same (the same notes in another ordering would form a different tune), the original order of a certain abductive procedure matters to its plausibility. Rearrange the order, and a different abductive procedure is formed. This temporality to abductive inquiry shall remain a presumption for the rest of the procedures discussed in this article.

This abduction fallacy concludes with some candidate ‘A’ for credibility, though that candidate must be deemed logically unacceptable here. The phrase “Therefore, A” fits well with deduction, and by convention, to induction. Instead of using ‘therefore’ with abduction, we shall use ‘so’ to indicate only an intended linkage between premises and tentative conclusion. Hence, “So, A” can only mean something like, “So, it appears that A is plausible,” or “So, perhaps A is believable.” In the realm of abduction, “So, A” is entirely compatible with “You shouldn’t regard A to be credible, since …” Where abduction is involved, inferences retain their conjectural and fallible status to some degree or another.

There are two intuitive reasons why this simple ‘affirming the consequent’ is a fallacy. First, almost no credibility can be given to an explanation when innumerable equally explanatory options (B, C, D, etc.) are available, since they haven’t been ruled out by Q or any other considerations. Call this the “Explanatory Plenitude Problem”. For example, if I blame a roaming raccoon for that sound heard outside my window late at night, I have arbitrarily picked one of many possible causes for that sound. Second, almost no credibility can be given to an explanation when simply positing some imagined A to be responsible for Q supplies no conception of the relationship between A and Q to make responsibility plausible. (And asserting “A is responsible!” is not a conception of the alleged relationship.) Think about it: why would a raccoon be causing such a noise? Call this the “Explanatory Responsibility Problem”.

Greatest confidence in A would be earned if we could arrive at: Only if A then Q, and Q, so A. Asserting “Only If A then Q” is an assertion both that no other B, C, etc. can explain Q, and the absence of other explanations is due to the way that the conceived ground of the relationship between A and Q that makes A responsible for Q is so concrete and compelling that A of necessity is responsible for Q. Then A would truly be the best explanation. Of course, that result is gained by effectively replacing abduction with deduction. The distance between simple abduction and straightforward deduction is vast indeed. How can that gap be narrowed?

With simple abduction, so far, some B (or C or D etc.) could be imagined easily to imply Q just like A. We are very far from "Only if A then Q."

What if we prevented any other optional explanations from becoming conceivable?

Q

Only If A then Q [because no other B, C, D, etc. seem thinkable, or if thinkable, they don't feel relevant]

So, A ["Deducible abduction"]

This is the conceptual conversion of "If A then Q" to "Only if A then Q" by psychological means. If B, C, D (etc.) can be made to seem unthinkable and impossible, then A appears to be the forced preference. But this tactic is only about biasing the mind, not learning about reality. Adjusting or interfering with people's minds to cause mental poverty or ignorance in a group cannot be a reasonable way to rule out the existence of alternative possible explanations. Deduction is a dangerous method to apply to substantial matters without caution. Conceivability may be a sign of real possibility, but inconceivability should never be taken to be a sure sign of impossibility about material matters. Limiting your knowledge to only what you find currently conceivable will prevent further learning. Nevertheless, a principle that "Foreign explanations are inconceivable among us!"—let this be labeled as the "Social Inconceivability of Options" principle—often operates among social groups.

More brains thinking just like us can't really increase deserved credibility. Back to observations then, so we need more Qs.

Q1

If A then Qs

Q2, Q3, … [induction]

So, A ["Inducible abduction"]

A seems to reliably predict lots of Qs, and more Qs keep coming. If one is primed for spotting more Qs as they occur, A can seem so predictive. But what about some rival explanation B? If B, then Q, and Q1, Q2, …, So B! And what about C or D as well, explaining those Qs that have been already observed? When all that has been observed is a reliable pattern of Qs, no A or B (etc.) seems to really put to any explanatory test. Just taking a series of Qs to be good evidence for A leaves one prone to a simply confirmation bias. However, where equally successful 'explanations' can proliferate just by imagination, the credibility for any single explanation falls towards zero.

So far, we have combined simple abduction with deduction and induction, without making much progress towards discovering a procedure deserving credibility. However, we do see where deduction and induction can be combined with abduction. Those two combinations can dominate abduction so that the procedure really isn't abductive anymore, but instead primarily deductive or inductive in nature. The three basic forms are:

ABDUCTION, because it is accounting for a surprising fact which is doing most of the credibility work.

Q!

If A then Q

So, A

INDUCTION upon abduction, because it is the iteration which is doing most of the credibility work.

If A then Q, and Q

If A then R, and R

If A then T, and T

So, A

DEDUCTION upon abduction, because it is the definition of A which is doing most of the credibility work.

Q!

If A's definition is suitably changed, then Q

So, A

Returning to our analysis of Inducible abduction, what about rival explanation B? If B, then Q, and Q1, Q2, …, So B! And what about C or D as well, explaining those Qs that have been already observed? We need to look at more than Qs. Two primary options open up at this stage.

EITHER

Qs, and If A then Qs

Rs, and If A then Rs

Ss, and If A then Ss

Ts, and If A then Ts

…

So, A [abductive induction – "Interative Abduction" – a sequence of similar abductions of things]

OR

Both Qs and Rs have feature F1

If A then Qs and Rs would have F1 [after defining A to 'effect' that analogous F1 displayed by both Qs and Rs]

Both Qs and Rs have feature F2

If A then Qs and Rs would have F2 [A's definition also 'effects' that analogous F2 in both Qs and Rs]

…

So, A [abductive deduction – "Coduction" – an abduction of similar features in things]

Let's discuss Coduction first. Although it applies abduction, it really is a kind of deduction. Deducing similar phenomena, these analogous features in both Qs and Rs, from A's definition is actually doing the plausibility work. This A can 'effect'—can 'be responsible for'—a curious series of analogous features (they only need be similar/analogous, not identical features) in two otherwise different things. There can be great intuitive plausibility attached to an A which can account for why separate things would display analogous features. We suspect some hidden thing responsible for the similar features to otherwise separate matters. We rely heavily on this basic Coduction in our human world, for example. When we visit a neighboring house and compare that house with our own, we might notice how that house has the same-sized kitchen to the left of the dining room, which also open up to the right onto the same-sized living room just as ours does, and the stairs proceed from that room up to two bedrooms just like our own house has, and so on, we soon will be thinking that both houses probably were designed by the same architect.

The detection of similar features across different things is a core intellectual capacity, and 'coduction' points at the inherent plausibility awarded to an explanation able to be responsible for that detected correlation. However, two main problems arise to severely limit the reasonableness of Coduction. First, the Explanatory Plenitude Problem will persist, since rival explanations B, C, D (etc.) will also try to be responsible for the same Fs of Qs, Rs (etc.). That opens the door for the second main problem: as rival A, B, C (etc.) set off to account for more and more analogous features among Qs, Rs, Ss, and even more things, our cognitive capacity to 'detect' similarities across disparate things will get powerfully exercised. We are too good at this capacity, though. Cognitive biases again make their influence felt, especially in our tendency to attend closely to coincidences and perceive strong patterns where only weak ones really exist. We can find analogous features in about any two different things with enough imaginative creativity. Our efforts will go into detecting analogous (we imagine) features of things, and not into the proposed connecting relationships between explanation A and feature Fs. This Explanatory Relationship Problem, as we can label it, will only grow. However, Coduction does show powerful explanatory power, even if controlling that power is evidently crucial. We will re-engage with Coduction after some detailed explorations into Abduction.

Next, concerning Iterative Abduction, while it applies abduction it really is a kind of induction. Intuitive plausibility can attach to an A which can account for why separate things display their detected frequency patterns after those patterns are discovered. A simple example illustrates the degree of plausibility that Iterative Abduction deserves. When I am upset to find that the garbage can behind the house has been overturned and garbage is littered about, I imagine a raccoon getting a midnight snack. The next day, not only has the garbage can been overturned and pillaged again, my neighbor's garbage can has also pillaged. After several similar incidents, I can't help but think that a raccoon has found a congenial picnic location.

There is a degenerate form of Iterative Abduction, where A is used to repeatedly 'explain' a series of features to Qs:

Qs have F1, and If A then Qs have F1

Qs have F2, and If A then Qs have F2

Qs have F3, and If A then Qs have F3

…

So, A [limited abductive induction – "Singular Interative Abduction"]

This narrower form can helpfully focus attention on a plausible explanation, but its explanatory power is severely limited. To continue my earlier example, I might notice how only certain kinds of food wastes left in my garbage appear to be consumed each night—just waste from foods containing nuts, peanut butter, or seeds. Every time that my garbage is invaded, I notice how either nuts, seeds, or peanut butter products appear targeted, so my suspicions turn towards a squirrel instead of a raccoon.

For both Iterative Abduction and Singular Iterative Abduction, some alternative explanation B could keep pace with similarly explaining the features of many Qs, or the series of Qs, Rs, Ss, etc., just as well as A. There is an additional risk that as more explanations C, D (etc.) also try to keep up, they become explanatorily empty. However, Singular Iterative Abduction in the long run, if perpetually successful, may arouse the suggestion that A and Qs may not be separate matters. If every significant feature of all Qs is 'reliably' effectuated by A, and A does not possess any of its own capacities not busily effectuating Qs, then the conceptual distinction between A and Qs fades. Perhaps Qs simply are manifestations of A from various 'perspectives'. For example, I still recall my astonishment as a child upon being told that traveling through fog is just like traveling through a cloud. Later, I learned why: there is hardly a difference between fog and cloud except altitude; one could fairly say that fog is just a cloud down upon the ground.

The "Principle of Identity of Effects" serves as a label for the proposal that where an explanatory thing always effects the same phenomena and never effects anything else, then those phenomena are just manifestations of that explaining thing. This principle can be very useful, but it must be applied cautiously, as discussions of more complex procedures shall illustrate.

Why wait to see what kind of A can keep predicting each and every thing that comes along? Perhaps we can define A more carefully up front. What if A can effect … everything!

[everything – all observed 'Zs' where a Z could be anything]

There is no Z such that If A then not-Z [by defining A just right and then deducing this second premise]

So, A [extreme deducible abduction – "Panoptical abduction" – abduction by everything observed]

While uncommon, this extremely imaginative sort of 'explanation' isn't alien to human thinking. When people long ago mostly lived in isolated villages, rooted to their local agricultural life, childish questions asking why the sky displays its bright lights, or why the landscape has its peculiar features, or why the people do the daily tasks they do, might (depending on local tradition) all be answered with ancient lore about a single high god who turns out to always be responsible for arranging all matters. Seeing "the hand of god" in all things remains an explanatory tactic available to theology to this day. Yet this tactic remains vulnerable to local ignorance; its plausibility relates to the "Social Inconceivability of Options" principle often operating among social groups. That vulnerability is exposed when one village discovers how the neighboring village credits everything to a different deity. That's the risk to crediting a lone A for all that explanatory work—what about some imaginatively defined B that can explain everything too? Deduction is needed again. What we need is an additional principle to add to the deduction process.

There is no Z such that If A then not-Z [deduced from A's definition]

There is no Z such that If B then not-Z [deduced from B's definition]

[everything]

If X is responsible for a set of things and Y is responsible for precisely that same set of things, then X = Y [Principle of Identity of Responsibles]

So, A ["Reductive panoptical abduction"]

The Principle of Identity of Responsibles has some intuitive power because one commonsensically doesn't expect some Z to really be entirely caused by both A and B, so only one is probably involved. For example, two neighboring villages, or two entire religions, may suspect that fewer than two deities are fully responsible for all creation (so they instead argue over the correct name for that singular supreme deity). However, that helpful intuition cannot logically identify which one, A or B, is actually responsible, or whether some unknown C might really be responsible. This Principle of Identity of Responsibles can't be generally valid. It only seems to be valid so long as there is nothing that could ever be unexplainable by A or B. Under those extraordinary conditions, we can't conceive of a difference that makes a difference. As Peirce judged, no logical difference remains between two hypotheses permanently having the same empirical consequences.[7] We can decide that the A/B distinction is just semantic, and we reduce them to each other so that only one explanation is really involved. In this atypical context alone, the Principle may be admitted.

But what about the way that it could still be the case that even "If A then [everything]", each particular Z never depends on A? Defining A "just right" to be logically compatible with all Z does not permit the inference that Z every actually depends on A. In fact, the vaguer A gets by definition to stay compatible with everything in the world, the less we are able to conceive of the grounds for a dependency relationship of any Z to A. A is assigned fewer and fewer traits and the remaining traits get more and more abstract. There is less and less in common between A and any particular thing, to the point where A shares almost nothing or nothing in common with things and cannot be understandably relatable to all things. (Theologians are familiar with the way that metaphysical conceptions of God can easily get vaguer the more that God is unlike creation.) The "Explanatory Relationship Problem" arises in the long run, in a new form. The claim that "A is responsible for each and every thing" can

[7] See Psillos [8, 135].

become explanatorily vacuous and the conception of A becomes empty. Label this as the “Explanatory Emptiness Problem.”

In order to avoid that explanatory dead end, we must return to a stage before deduction was allowed to tempt us to define A with excessive ‘explanatory’ power. We therefore return to this stage:

Qs, and If A then Qs

Rs, and If A then Rs

Ss, and If A then Ss

Ts, and If A then Ts

…

So, A [abductive induction – “Interative Abduction” – a sequence of similar abductions of things]

Yet it is still the case that some alternative B might keep up with explaining Qs, Rs, Ss, etc. That possibility of competition should not get ruled out. So we must restrain our conception of A in advance.

3 Abduction Controls

Let’s try to control the definition of A so that it only has a delimited amount of traits and powers.

Qs, Rs, Ss, and Ts!

If A then Qs

If A then Rs

If A then Ss

If A then Ts [and given A’s definition, by deduction we see that there are no more things for A to explain]

So, A [deduced abductive induction]

All the same, we won’t wait long for some B, C, and D to show up to explain Qs, Rs, Ss, and Ts too. It’s too easy to conceive of some new B (etc.) such that E ‘explains’ a given list of Qs, Rs, Ss, and Ts already observed. (A cat or a dog, rather than a raccoon, may be getting into each house’s garbage cans on my street.) If many conceivable causes for the same observed phenomena can be considered, what can be called the “Explanatory Plenitude Problem” arises to diminish confidence in any of the possible A, B, C, D, etc.

Delimiting the conception of A up front was too hasty. We must limit the explanatory responsibilities of A without delimiting them too much up front. Our answer is this: we shall permit A (and B, etc.) to be defined generously up front, permitting it to potentially be responsible for matters not yet observed.

Qs!

If A then Qs

Rs!

If A then Rs [given A's definition, by deduction we see how Rs would be expected from A]

So, A ["limited interative abduction"]

Of course, some alternative explanation B could also turn out to expect Rs too. What could throw the advantage to A again?

A's advantage would be due to greater explanatory reach. How many novel phenomena might A be able to 'predict' after their discovery—how much can A 'retrodict' in the long run?

Qs!

If A then Qs

Rs!

If A then Rs [given A's definition, by deduction we see how Rs would be expected from A]

Ss!

If A then Ss [given A's definition, by deduction we see how Ss would be expected from A]

…

So, A [deducibly abductive induction – "Retrodicted Abduction"]

This procedure can continue for a long time, depending on the initial definition of A and how many kinds of phenomena can be elicited from it. (If only a dog would knock over lots of garbage cans, and eat all of the food waste, and dig a hole nearby to bury a steak bone, etc., then suspicions turn towards blaming a dog.) There is a similar inductive version to Retrodicted abduction, "Retrodicted Induction":

Qs!

Suppose that If A then Qs [after designing A's definition quite vaguely, to expect Qs along with plenty of other unspecified matters]

Rs!

Suppose that If A then Rs [now expecting Rs from A's vague definition too]

Ss!

Suppose that If A then Ss [now expecting Ss from A's vague definition too]

…

So, A [abducibly deductive induction – "Retrodicted Induction"]

Retrodicted Induction superficially looks like an abductive procedure. It surely is far more suspicious, because A's definition is designed in advance to 'explain' not just some initial Qs but also plenty of other vaguely indicated matters, so that any chosen Rs, Ss, and Ts (etc.) can get 'explained' when they show up later. (If my partner gets fearfully convinced that a bear is roaming the neighborhood, without knowing much about bears, that suspicion gets stronger every day that a mess is discovered, because "That's apparently just what a bear would do!") Retrodicted Abduction seems less suspicious by comparison, because at least A has the modestly greater merit of not being pre-designed to vaguely fit with some selected set of phenomena observed later on.

For an explanation A supported by either procedure, some rival explanations B, C, and D will try to keep pace, but a failure rate will build up among them. It is possible that some E will outpace all other explanations tried so far, by retrodicting more explained phenomena than the rest without exhausting its explanatory powers.

If two explanations, E and F, have explained all surprising Qs, Rs, …, Zs without signs of explanatory exhaustion, perhaps E and F are really about the same thing? Apply the Principle of Identity of Responsibles here: If X is responsible for a set of things and Y is responsible for precisely that same set of things, then X = Y. However, that Principle is not valid in this context, where it is not known whether their sets of explained phenomena will remain the same, so therefore E and F cannot be known to be identical to each other.

Furthermore, conceptions of E and F become vaguer as they repeatedly come up with post hoc ways to 'explain' what gets observed. By the Principle of Explanatory Emptiness, explanations E and F (etc.) risk becoming vacuous as they race each other to maximize phenomena explained. It seems to be a dead end to expect maximal retrodictions from explanations.

More pressure must be put on the explanations A, B, C etc. by the things Q, R, S (etc.) that they are supposed to be responsible for. Two primary options emerge and diverge at this stage.

The "Predicted Independent Phenomena" scenario, in which the definition of A gets induced to expect an iteration of Qs, Rs, etc.:

If A then Qs [given A's definition, by deduction we see how Qs would be expected from A]

A pattern of Qs gets discovered!

If A then Rs [given A's definition, by deduction we see how Rs would be expected from A]

A pattern of Rs gets discovered!

…

So, A [deduced inductive abduction – "Predicted Abduction"]

OR

The "Predicted Analogous Phenomena" scenario, in which the definition of A is used to deduce features, and a iterated series of Fs are predicted for Qs and Rs:

If A then Qs and Rs have F1 [from A's definition, by deduction feature F1 is expected in both Qs and Rs]

Qs and Rs have F1!

If A then Qs and Rs have F2 [from A's definition, by deduction feature F2 is expected in both Qs and Rs]

Qs and Rs have F2!

…

So, A [induced deductive abduction – "Predicted Coduction"]

Predicted Abduction is the first procedure in this development which genuinely deserves some credibility. It combines two important features. First, the definition of A remains stable throughout the iterations of abductions (this will become highly significant), so iterations are deduced from the same definition to A. Second, the definition of A is applied to make "novel" predictions about what would be discovered, and those discoveries are made. This is abductive risk, which should not get lost. (Knowing more about dogs than bears, I can predict further signs of canine scavenging around the neighborhood, and watch for confirmations.) So long as the iterated pattern of successful predictions can continue, A can enjoy some deserved credibility. How much credibility? Well, we mustn't get too attached to A, because its run of explanatory luck may not be long, especially because the definition of A is inflexibly rigid. How much can A really keep predicting?

We consider Predicted Coduction next. Like its simpler version already considered, this procedure must eventually suffer from the Explanatory Relationship Problem. Recalling my neighbor's house, sharing striking features to its floor plan with my own house, we might notice further similarities as we get more convinced that one architect designed both houses, while overlooking major dissimilarities and ignoring the possibility that two different architects coincidentally designed similar plans. All the same, an explanation A using Predicted Conduction will resist its dismissal, on the grounds that it has impressive explanatory power, by accounting for surprising similar features to different matters. That impressive ability to focus attention only on similarities, in the long run, is exactly what should erode its credibility upon reflection.

As A is applied for finding expanding analogous features to separate phenomena, great pressure will compel the conception of A to change and develop, so that it can 'effect' additional features to new things while still accounting for already explained features.

4 Abduction Inflation

If Predicted Conduction is applied in a more flexible manner, then it will actually look like this in practice:

The "Predictably Analogous Phenomena" procedure:

If A then Qs have features F1 [given A's definition, by deducing how Qs having F1 are expected]

Qs have F1!

If A then Rs also have analogous features F2 [after adjusting A's definition, then deducing how Rs having F2 are expected, while still deducing Qs with F1 too]

Rs have F2!

…

So, A [inducibly deductive abduction – "Predictable Coduction"]

This flexible procedure of Predictable Coduction deserves its name, because after a while this mode of explaining gets somewhat predictable. Primed by an initial supposition of A, our cognitive capacities search out novel features of curious matters, and then we ponder adjusting our conception of A just right to 'predict' some analogous feature to those matters getting explored. For example, suppose my friend makes his first visit to a foreign country and barely understands its language. Overhearing native speakers warn against the ‘zumzum’ (names have been changed to protect the innocent species), my friend infers that this zumzum is a nasty thing to watch out for. That night, noises against the window disturb her rest, she notices a red bite mark on her ankle the next day, and on the third day she catches a glimpse of something small crawling on her sandwich. She now thinks that a single insect is behind these manifestions: this ‘zumzum’ is a flying bug, that seems to also be a bug that flies onto people to bite them, as well as a hungry biting bug that lands on food. Confident that she has identified what this zumzum is, after blaming it for a series of incidents “bugging” her, she would be surprised to learn from local inhabitants that ‘zumzum’ doesn’t actually refer to any of those things.

The abduction aspect to Predictable Coduction inflates the conception of A to keep up with whatever evidence is brought before it. We simultaneously ‘guess’ at what A should be conceived as while we try to detect just those suitable analogous features which will ‘fit’ well enough with a revised conception of A. After those features are ‘found’, then we ‘confirm’ that revised conception to A, and even higher credibility gets assigned to A. There is no lack of cleverness to this procedure, since it may be difficult to simultaneously imagine a modified A that can ‘effect’ just the right analogous features to some new matter still getting explored. This sort of procedure can be productive at the ‘cutting edge’ of new knowledge, where features of poorly understood things are not clear at all. That’s why the ‘detection’ of ‘predicted’ features can happen more easily, and ‘confirmations’ to A are more frequently available than with any well-established subject matter. This also explains why, despite the difficulties to ‘predicting’ novel features to things, rival explanations B, C (etc.) can manage to compete and survive alongside A.

As suspicious as this flexibility to A (and B, C, etc.) must be, Predictable Coduction marks a needed transition in the development of explanatory inference, precisely because A’s conception is changing to respond to more and more evidence. Not only is the conception of explanation A central to the evaluation of these inferences, a feature of any material inference, any potential growth for a material inference’s explanatory power lies in the deliberate inquiry-led modification to explanations. Down that road lies the full explanatory power of material inferences harnessed to inquiry-driven evidence gathering.[8]

Although Predictable Coduction displays some explanatory potential, a degenerate form of Predictable Coduction refuses to make specific predictions about what phenomena would be observed. Instead, this procedure first notices unusual features to newly discovered things, and then adjusts its explanation A so that those features ‘fit’ a pattern of analogy with previously explained features to other phenomena.

The “Deducibly Analogous Phenomena” procedure:

Qs have F1!

If A then Qs have features F1 [given A’s definition, by deducing how Qs having F1 are expected]

Rs have F2!

[8]Consult Brigandt [21].

If A then Rs also have analogous features F2 [after adjusting A's definition, then deducing how Rs having F2 are expected, while still deducing Qs with F1 too]

Ss have F3!

If A then Ss also have analogous features F3 [after adjusting A's definition again, then deducing how Ss having F3 are expected, while still deducing Rs with F2 and Qs with F1 too]

…

So, A [deducibly abductive induction – "Iterative Coduction"]

Iterative Coduction can attract even more credulous believers than Predictable Coduction, because this procedure can be repeated and applied to almost anything novel and somewhat mysterious, making it appear that A is endlessly 'explaining' many new curious matters and yielding explanatory connections among them. As A conceptually inflates, its believability seems to grow. This attractive credulity is the reason why fascination with the "uncanny" linked with oft-repeated superstition is prevalent across human societies. For example, faulting malevolent spirits for all sorts of human miseries is not a common custom of only pre-modern times. (Societies accustomed to such 'explanations' typically host 'experts' predicting and negotiating with these hidden powers.) Iterative Coduction is a procedure utilized by mythic thinking, because it is both scalable and hierarchical. It can be broadened endlessly across ranges of phenomena, and higher-level explanations can superstructurally be "Coduced" to unify important features displayed by lower level explanations.

Claude Lévi-Strauss located the divide between scientific and nonscientific mentality here. Both care for evidence and explanation, but mythic mentality spins its intense practical obsession with everything in the environment into all-encompassing explanatory webs.[9] It is unnecessary to appeal to Lévi-Strauss's controversial theses; a reliance of mythic cosmology on Iterative Coduction is evident. When a single and supremely explanatory web uniting all explanatory webs is creatively developed, the realm of mythic cosmogony emerges, perhaps including agent deities or at least supreme powers. Where a grand mythic web of explanation is sustained by structured inculcation across generations, it acquires features commonly associated with religion.

The second and third intellectual strategies both lead to the same desired result, a guarantee that A will be the 'best' and final explanation, amounting to a procedure which we label as Deduced Coduction.

Qs have F1, and If A then Qs have F1

Rs have F2, and If modified-A then Rs have F2

Ss have F3, and If modified-A then Ss have F3

…

Principle of Explanatory Fertility/Principle of Identity of Responsibles

So, A [inducibly abductive deduction - Deduced Coduction, by pan-fertility or by pan-responsibility]

---

[9] Lévi-Strauss [22].

In addition, recalling the procedure of Panoptical Abduction, a third strategy to 'guarantee' that A is the best and final explanation can be rationalized. This third way is called "Reductive Pancosmism" because everything that can ever be in the cosmos is reduced to an effect of a single ultimate explanation.

There is no Z such that If A then not-Z [deducible from A's flexible definition]

[everything]

Principle of Explanatory Fertility/Principle of Identity of Responsibles

So, A [abducible inductive deduction, Reduced Pancosmism, by pan-fertility or by pan-responsibility]

Metaphysical systems, some idealistic and others materialistic, can resort to Deduced Coduction and Reduced Pancosmism, as the history of philosophy displays. Core differences among kinds of theologies are also traceable back to these procedures.

Materialisms tend to prefer Deduced Coduction while idealisms typically rely on Reduced Pancosmism. Four primary types of 'mono-theology' also distinguish themselves here. The development from magical and superstitious imagination towards mythic ideas about hidden agents and guiding powers reaches its culmination in cosmogonic religions and rationalizing theologies. Hinduism's Upanishads and Advaita Vedanta relied on Reduced Pancosmism by pan-fertility, as the transcendent reality endlessly generates (perceptibly or imperceptibly) all dualities and natural entities. Christianity's monotheism relied on Reduced Pancosmism by pan-responsibility, as the lone Creator effected (directly or indirectly) each natural thing. Greek polytheism produced a theology that can be called "panpolytheism" which declared that other nations' pantheons apply their local names to the one true set of gods genuinely responsible for all events, hence relying on Deduced Coduction by pan-responsibility. Roman Stoicism relied on Deduced Coduction by pan-fertility, by attributing to uniquely supreme powers ('deities') the responsibilities for harmonizing all of nature's generative cycles and supportive habitats.

5 Abduction Development

Let us return back to two procedures already enumerated, positioned before mythic thinking branched away: the Predicted Independent Phenomena scenario and the Predictably Coductive Abduction procedure. To recall:

The "Predicted Independent Phenomena" scenario:

If A then Qs [given A's definition, by deduction we see how Qs would be expected from A]

A pattern of Qs gets discovered!

If A then Rs [given A's definition, by deduction we see how Rs would be expected from A]

A pattern of Rs gets discovered!

…

So, A [deduced inductive abduction – "Predicted Abduction"]

AND

The "Predictably Analogous Phenomena" procedure:

If A then Qs have features F1 [given A's definition, by deducing how Qs having F1 are expected]

Qs have F1!

If A then Rs also have analogous features F2 [after adjusting A's definition, then deducing how Rs having F2 are expected, while still deducing Qs with F1 too]

Rs have F2!

…

So, A [induced deductive abduction – "Predictable Coduction"]

Predictable Coduction enjoys an enormous advantage over Predicted Abduction: its definition of A is permitted to developmentally change. Turning to an example from physiology, the heart was long ago connected to the flow of red blood. If the heart by definition rhythmically puts red blood out through the arteries for consumption by the body, then further events would be observed; the ancient Greek physician Galen noted pulsing red blood from cut arteries of the limbs, and hearts pumping blood from its chambers during vivisection. Galen's delimited definition for the heart allowed centuries of physicians to ponder how the heart makes red blood, why heart valves restrict blood flow direction, where blue blood comes from, why some arteries are conveying blue blood, and many more mysteries. Harvey's seventeenth century discoveries about the heart and circulatory systems were predicated on flexibly defining the heart differently. The heart does not make blood and blood isn't consumed but only transformed, as the blood is pumped out through arteries, back towards the heart through veins, and cooled by the air in the lungs during a side trip (he was unprepared for reconceiving the lungs).[10]

Admirers of Predicted Abduction might point to its definitional rigidity to endorse its higher credibility, since A doesn't receive post hoc modifications suspiciously capable of predicting new phenomena. It is the case that Predictable Coduction's flexibility exposes this procedure to the high risk of degenerating into post hoc pseudo-explanations, explanatory vacuousness, and superstitious thinking. However, puritanical admiration for Predicted Abduction would be misguided. After all, the definition of A just to explain the initial Qs had to be forged from available prior notions, and A's conception is almost always some modified older idea. Furthermore, revolutionary theories in the history of science always undergo modification and development as they are extended to wider and disparate phenomena. It is very difficult to identify some useful hypothesis which never changed at all from its initial conception to its fullest utility. Nor does theoretical rigidity serve as a reliable predictor of scientific success—many of best confirmed scientific theories underwent dramatic development in the course of their thorough testing. Indeed, theoretical rigidity seems to always part of the regrettable story to discredited explanations.

[10]See Shackelford [23], chap. 2.

Let us honestly admit that explanatory rigidity is no safe path towards credibility. Permitting explanatory flexibility, we next generate the procedure of Predictable Abduction:

The "Predictable Independent Phenomena" procedure:

If A then Qs [given A's definition, by deduction we see how Qs would be expected from A]

A pattern of Qs gets discovered!

If A then Rs [after adjusting A's definition, then deducing how Rs would be expected from A, while still deducing Qs too]

A pattern of Rs gets discovered!

…

So, A [deducibly inductive abduction – "Predictable Abduction"]

At this stage, with two tentatively viable procedures, we can see how far we have come all the way from simple abduction to inducibly deductive abduction and deducibly inductive abduction.

Both procedures flexibly alter an explanation in the course of anticipating and predicting novel patterns to things or novel features to things. As we discussed, Predictable Coduction is a dangerous procedure to explore. Unless strict controls are placed upon modifying A in the process, A soon enjoys vast explanatory fertility at the cost of becoming explanatorily vacuous and/or A only 'explains' things as they get discovered, generating mythos. All the same, discerning an explanation behind the prediction of analogous features to otherwise separate things feels familiar to us and possesses an undeniably powerful cognitive appeal. Things really are more deeply connected that they may appear, as a natural matter. The tougher question is how to develop this Coductive procedure in order to weed out poor explanations from better ones. Let's set this issue aside, for now.

Predictable Abduction also look familiar and compelling, for obvious reasons. Let's try to develop this Abductive procedure in order to weed out poor explanations from better ones. We should first admit that good explanations try to explain already curious matters. We should also admit that if we will be permitting conceptions of explanations to get set down and also modified, those explanatory features should have the capacity (somehow) to effectuate Qs, Rs, and so on, so that there is some conceivability to A (A at least has its set of capacities C1, C2, etc.), and also some conceivability to the basis for the effective relationship between the explanation and the phenomena getting explained.

This next procedure controls the set of capacities so that they are immediately put to use to explain how Qs, Rs etc. get effectuated.

This AIA procedure: A and its capacities Cs are abductively related, and then that relation is induced to effect Qs, Rs, etc.

Qs!

(If A then C1), then Qs would be effected from C1

(If A then C2 too), then Rs would be effected from C2

Rs!

(If A then C3 too), then Ss would be effected from C3

Ss!

…

So, A(Cs) [abductively inductive abduction - AIA]

The problem with AIA is that crediting A with multiple capacities, accumulating to account for more and more phenomena, leave the conception of A with a set of otherwise unrelated and ad hoc capacities. Looking to the history of science, the example of germ theory illustrates AIA's potential and limits. The hypothesis that tiny living forms or 'germs' are involved with sepsis and pestilence traces back to the seventeenth century. Over subsequent centuries, germs were occasionally raised by speculative physicians and botanists to account for very different phenomena from infections, boils, and fevers to plagues among humans and livestock, and even to afflictions to plants and crops. The concept of 'germ' was left far too vague, and related to varying phenomena in an ad hoc manner, leaving little solid information for researchers to work with. By the late nineteenth century, biologists were distinguishing bacteria, molds, fungi, worms and other parasites, and many more kinds of microorganisms, permitting scientists to link specific pathogens with certain maladies. AIA by itself credits A with various capacities 'generated' from A, but not because we really understand A, but only because they would conveniently produce the phenomena. There's no reason given why A would have these Cs, or how they relate to each other. Instead of the Explanatory Vacuousness problem, there is a mysterious Capacity Overload problem. Furthermore, any rival explanation B, C (etc.) can keep pace by including those capacities too, so no explanation can really gain any advantage over sufficiently imaginative rivals even in the long run.

To avoid the appearance of arbitrariness where Cs are gradually elicited from the conception of A, it is possible to first list all the Cs that the explanation should have, and then check to see if those capacities do effectuate further Rs, Ss, etc.

This IAA procedure: A and its capacities Cs are inducibly related, and then that relation is abductively effectuating Qs, Rs, etc.

C1, C2, C3 … imply A which has these Cs, and Qs would be effected from C1

Qs would be effected from C1

Qs!

Rs would be effected from C2

Rs!

Ss would be effected from C3

Ss!

…

So, A(Cs) [induced abductive abduction - IAA]

The problem with IAA is that the initial list of capacities for A would have to be amazingly predictive of not just the initial Qs, but also the Rs, Ss (etc.) in advance. It would require the most extraordinary guessing at just the right needed capacities to accomplish this procedure successfully for very long. If that amazing guessing feels somehow believable to someone, that person would find A extremely credible. But more skepticism is recommended. This procedure can work well on someone ignorant about Qs, Rs, Ss, etc., so that the fraudulent claim is made that this explanation A arose long ago before all these Zs were discovered but A still managed to 'predict' them. For example, the attraction to "Ancient Wisdom" that amazingly anticipates today's matters and recently discovered phenomena falls into this category of intellectual fraud. Of course, that explanation A was actually invented recently after all those Zs had been discovered, and so this procedure effectively collapses into IAD, Deduced Coduction, which is a 'theological' procedure.

A similarly suspicious procedure, IIA, simply assigns a set of Cs to an A which are capably of effecting any number of Zs of some general character.

This IIA procedure: A and its capacities Cs are inducibly related, and then that relation is inductively/effectuating Qs, Rs, etc.

C1, C2, C3 … imply A which has these Cs, so that various Zs would be effected from one or another of the Cs

Qs!

Rs!

Ss!

…

So, A(Cs) [induced inductive abduction – Elicited Abduction]

The way that Elicited Abduction won't say in advance much about what specific sorts of phenomena these Zs will be must arouse suspicion and skepticism. When specific Rs, Ss (etc.) get detected and elicited into service, it would be too easy to say that those are the among the Zs 'predicted' by A having those Cs. For example, a pseudo-scientific theory such as astrology relies on this Elicited Abduction procedure, in which the capacities of heavenly bodies are supposed laid down by the theory, and they in turn are responsible for vague sorts of elicited Zs noteworthy here on earth. When interesting things do happen on earth, such as Rs and Ss (you are born with a certain temperament, or specific things happen to you today) then the astrologer announces that those Rs and Ss were indeed among the Zs "foretold" in the heavens. This procedure is most plausible to people who are already quite susceptible to confirmation biases.[11]

Instead of specifying A's Cs in advance, we could return to a method that adjusts the Cs possessed by A gradually, as the procedure goes along from prediction to prediction.

The ADA procedure: A and its capacities Cs are abductively related, and then that relation is deductively applied to imply analogous features to Qs, Rs, etc.

Qs have F1!

Supposing (If A then C1), then Qs would have feature F1

Supposing (If A then C2 too), then Rs would have feature F2

Rs have F2!

[11] On pseudo-science in general, the reader may begin by consulting Pigliucci and Boudry [24].

Supposing (If A then C3 too), then Ss would have feature F3

Ss have F3!

…

So, A(Cn) [abductively deductive abduction – ADA – Abductive Coduction]

Abductive Coduction manifests a tendency, also seen in IIA, towards a strong resistance to any disconfirmation. Suppose that after supposing that A has C4 and expecting some Ts with feature F4, those Ts aren't showing up as predicted. Has A suffered from a disconfirmation, so that doubt instead of credulity is earned here? No, the proponent of A will explain, all that has happened is that the fallible abduction that A has C4 was hasty and mistaken. A hasn't been disconfirmed at all—only C4. In fact, this bad prediction has yielded credible information about A, that it lacks C4. But A's existence remains a secure matter, we will be assured. Over time, by this selective procedure, A will acquires capacities (let's say) C1, C2, C3, C6, C12, C15, C19, C23, and C37. Look at how many capacities of A have been "confirmed"!

Although Abductive Coduction is minimally proto-scientific, since it is at least imaginatively experimental, reliance on this procedure would be unwise. The danger is that devout conviction that A is real can be sustained in foolishly credulous people for a long time by applying this hit-or-miss method. Conspiracy theorists rely on selective Abduction Coduction; they imagine that important events are really the outcomes of plots by a secretive organization, let's say. Which events? That's the puzzle-solving fun to being a conspiracy theorist—only the truly significant events would occupy such a powerful and secret organization, so one must weed through each year's worth of notable events to discern just the ones that could and would be accomplished by this secret organization (using their money? their threats? their political machinations? their overseas support? and so on) in a timely and effective manner.

The logical fact remains that A would not be seriously tested by Abductive Coduction, although whoever is assigning Cs to A and garnering some confirmations would be a very good guesser. Good guessing could also be displayed by proponents of a rival explanation B, C (etc.), as well. There could be something real about A and its confirmed capacities, but this is a poor procedure for credibly figuring out what is really the best explanation. There are fewer fruitless debates than those between adherents of rival conspiracy theories.

Abductive Coduction is hence susceptible to degeneration into hasty judgment, cognitive bias, fallacious inference, and even outright trickery. It is the method preferred by a fraud who might fool spectators into thinking that he or she possesses extraordinary powers. With enough imagined 'capacities' for making forecasts, diagnosing illnesses, reading others' thoughts, doing impossible feats, and so on, the busily risk-taking fraud can luckily (or skillfully, with some magic tricks) accomplish some unexpected results once in a while before surprised onlookers. When the credulous people in the crowd have "seen with their own eyes" just a handful of 'successful' confirmations to this fraud's amazing capacities, they don't attend as much to the disconfirmations. This risk of degeneration into the "Fraudulent Powers" problem leaves Abductive Coduction in generally poor repute.

6 Abduction Evolution

We proceed to a more complex stage, for working out procedures that exercise stricter controls on the capacities assigned to the conception of A. The next procedure in the sequence is DAA.

The DAA procedure: A's capacities are deduced from A's definition, but then they are abductively related to Qs, Rs, Ss (etc.)

Qs!

Suppose (only if A has C1), then Qs

Suppose (only if A has C1-2), then Qs & Rs

Rs!

Suppose (only if A has C1-3), then Qs, Rs & Ss

Ss!

…

So, A(Cn) [deducibly abductive abduction – "Strict Abduction"]

Unlike the simpler suspicious procedures AIA, IAA, IIA, and ADA, which run into their troubles by not strictly controlling the capacities assigned to A's definition, DAA exercises very strict control over modestly modifying the conception of A. Only the capacities required to account for the phenomena are attributed to A, and whatever the definition of A may be, that definition is only permitted to be compatible with those Cs applied in the procedure. No other conceptions of A, beyond those Cs proposed to account for Qs, Rs, Ss (etc.) are regarded as relevant. DAA has similarities with the simpler procedure of Predictable Abduction. However, instead of allowing the definition of A to be as broad as desired and adjusting it whenever it is convenient to predict some Rs, Ss (etc.), as Predictable Abduction allows, DAA does not permit the definition of A to range beyond whatever is minimally necessary for it to have its explanatory capacities. That is why we may label DAA as Strict Abduction.

Strict Abduction has five additional merits. First, whenever it being used, any particular time the conception of A has only one clear definition and set of capacities. Second, due to this bounded clarity, a community of inquirers can apply A together and everyone can agree upon what the explanation is and what it so far entails. Third, although a community will disagree over what new capacities A should have for increasing its predictive range, both the current definition of A and the presently assigned capacities place compatibility constraints on the sort of new capacities that can be assigned to A. Fourth, if a new prediction goes badly, only the relevant implicated capacity of A must be doubted, and not the rest of the capacities of A, preserving what explanatory power A had already earned. Fifth, the expansion of A's capacities and its explanatory range can halt whenever the community finds no work for A to do presently, but A can be put to work again in the future.

Comets can illustrate Strict Abduction. During the late 1500s, astronomer Tycho Brahe's observations suggested that comets are celestial (not atmospheric) bodies due to their observed trajectories; if celestial, they would be distant from the earth, and Brahe's parallax measurements indeed indicated their immense distance and vast size. By 1604 Johannes Kepler added that the sun's rays cause a comet's head to expel a stream of nebular material shining by the sun's light; his idea fit well with the usually overlooked way that a comet's tail always points away from the sun. This celestial, naturalistic, and causal explanation for comets hasn't essentially changed, but only supplemented. If comets journey between the planets, their paths must also be affected by the sun. By the late 1600s, Isaac Newton determined that a comet approaches the sun, swings around behind it, and departs away from the sun, and he explained why a parabolic path due to gravity would be typical for many comets. Also, Newton suggested that the sun would heat a close comet to incredible temperatures, so the head of a comet must be dense while the tail would be vaporous. Later investigations confirmed these hypotheses, completing the basic theory of comets.[12]

The transition from Predictable Abduction to Strict Abduction marks the boundary into scientific reasoning. Predictable Abduction, Predictable Coduction, and even Abductive Coduction are proto-scientific. They also can be put to use for pseudo-scientific and theological ends, as the proto-scientific is simultaneously logical, mythological, theological, and scientific. All four procedures are cohabitants of a broad realm of "speculative" thinking, or what the ancient Greeks called "inquiry into nature" (not excepting the cosmic gods), which is an arrival place of many simpler methods and a departure point for complex procedures going in different directions. Several civilizations arrived at this generative nexus of the proto-scientific and proto-theological.[13]

Only Strict Abduction ventures on into fully scientific methodology. That journey leaves behind preferences for vaguely conceived yet richly imagined explanations that elicit credulity by appealing to familiar notions, cognitive biases, and selected evidence than genuine predictive power. A scientific hypothesis restricts the capacities (properties, powers, etc.) of a hypothesized thing to some fairly delimited set, and those capacities are stable and habitual. The logic of testing hypotheses requires such features; specific predictions must be made and confirmed, so postulated entities must behave in patterned ways under specified conditions. That is why science has an innate preference for proposing constant impersonal capacities to explain observed regularities and mundane matters, leaving mythic and religious thinking to imagine less than predictable (fickle and willful) agents to account for singular extraordinary events.

A close variant to Strict Abduction is DDA—Deducible Coduction—in which A and its capacities Cs are deducibly related, and then that relation is deductively applied to imply analogous features to Qs and Rs.

DDA:

Qs and Rs have F1!

[12] Heidarzadeh [25], chap. 4.

[13] On that Greek nexus, consult Buxton [26], Morgan [27], Wians [28], Mikalson [29].

Only if A then C1, then Qs and Rs would have feature F1

Only if A then C2, then Qs and Rs would have feature F2

Qs and Rs have F2!

Only if A then C3, then Qs and Rs would have feature F3

Qs and Rs have F3!

…

So, A(Cn) [deducibly deductive abduction – DDA – Deducible Coduction]

Deducible Coduction is also a basic, but soundly scientific procedure. It utility is limited to the investigation of two different kinds of things which share in many common features. Recalling how fog banks are practically low clouds, their common manner of refracting and obscuring light (F1) is due to their composition of tiny water droplets (A). With enough water particles suspended in the air (C1), both clouds and fog banks would obscure light in their characteristic way. Water particles condense from water vapor (C2) when just a few degrees separate the air temperature and the dew point, so both clouds and fog would form when those conditions prevail (F2), regardless of altitude (although wind matters). Further properties of condensed water vapor account for additional common features to both clouds and fog.

In the long run, Deducible Coduction can permit a long iteration of successful predictions that Qs and Rs share in every significant feature. If there seems to be no significant feature that Qs and Rs do not share and A's capacities have predicted all of them, a further inference seems plausible: the genuine connection between Qs, Rs, and A must be far tighter than originally postulated. Perhaps Qs and Rs are simply two ways for A to effectively manifest itself (so that A and Qs aren't really two separate matters, nor are A and Rs—e.g. fog *is* cloud *is* amassed water droplets). Alternatively, going even further, there really was no A in the first place because Qs and Rs really are the same thing understood from two different 'perspectives'.

The first suggestion amounts to a "Principle of Identity of Effectables" while the second suggestion amounts to a "Principle of Identity of Correlatables". The Identity of Effectables means that Qs and Rs are dual manifestations (or 'properties', etc.) of one single underlying A. The Identity of Correlatables means that there never really was any A, since it is now deemed explanatory eliminable, so that Qs and Rs were really the same thing all along. (Further inquiry could next determine if R has ontological priority so that Q is 'actually' just R, or the reverse). These two Principles would function in two different procedures as follows:

Application of the Principle of Identity of Effectables

Qs and Rs have F1!

Only if A then C1, then Qs and Rs would have feature F1

Only if A then C2, then Qs and Rs would have feature F2

Qs and Rs have F2!

…

For all significant Fs of Qs and Fs of Rs, each Fn of Qs = some Fn of Rs [by inductive searching and discovery – "Identity of Features"]

Principle of Identity of Effectables – Where all of A's capacities effectuate Qs & Rs Identity of Features, then A = Qs and A = Rs

So, A(Cn) = Qs & Rs [Maximal Coduction]

OR

Application of the Principle of Identity of Correlatables

Qs and Rs have F1!

Only if A then C1, then Qs and Rs would have feature F1

Only if A then C2, then Qs and Rs would have feature F2

Qs and Rs have F2!

…

For all significant Fs of Qs and Fs of Rs, each Fn of Qs = some Fn of Rs [by inductive searching and discovery – "Identity of Features"]

Principle of Identity of Correlatables – Where all of A's capacities effectuate Qs & Rs Identity of Features, then Qs = Rs

So, Qs = Rs [Maximal Reduction]

Only each scientific field of inquiry can be responsible for judging the circumstances and background knowledge that permit the application of either Maximal Coduction or Maximal Reduction. These are fallible applications under the best of circumstances, since the possibility of rival explanations doing an even better job of explaining Qs and Rs, or a different job of relating Qs and Rs to other phenomena, cannot be ruled out in advance. Regarding banks of fog as just low-lying clouds because their composition and conditions for formation are so similar is an illustration of Maximal Coduction. An illustration of Maximal Reduction is the fate of Lavoisier's 'caloric fluid', an elemental gaseous substance within all bodies which flows from hotter to cooler regions. Chemists solved many experimental problems using caloric theory, while pondering how caloric fluid would also be the basis for the kinetic motion of molecules responsible for temperature. By the mid-1800s, Rudolf Clausius and James Clerk Maxwell demonstrated that the transfer of heat is just the redistribution of molecular kinetic energy obeying the principle of conservation of energy, so 'caloric fluid' was discarded.

There is one more procedure to this stage, AAA, which combines the merits of Strict Abduction with those of Deducible Coduction.

The AAA procedure: A's capacities Cs are abductively proposed from A's prior explanatory successes, and then they are abductively related to Qs, Rs, Ss (etc.). Let W(1-n) and Y(1-n) stand for any related series of Qs, Rs, Ss, Ts…Ns. Also, we define W(1-n) and Y(1-n) as an "analogous series" where common features found among all members of W(1-n) are also found, in analogous form, among all of Y(1-n).

If (Only if A's Cs have predicted a series of unexpected W(1-n)) then (an analogous series of unexpected Y(1-n))

Y(1-n)!

So, A [abductively abductive abduction – "Productive Abduction"]

For Productive Abduction, the C's of A are a "model" applied to the impressive effectuation of one "structure"—a series of W(1−n)—and that model additionally permits the successful prediction of another analogous structure of Y(1−n). To illustrate Productive Abduction, consider the development of cell theory in biology. Seeking the fundamental basis of life, the idea of a cell having its own cell wall and internal organic processes led botanists towards confirmations from studying microorganisms and plants. By the 1830s, this model was successfully applied to animal tissues, where cells displayed a similar construction and physiological functionings, and the cell was confirmed as the basic organic unit for all life forms.

Explanations confirmed by Productive Abduction deserve credibility. This procedure exploring the explanatory productivity of models is respectably scientific, while remaining naturally fallible.

7 Abduction Heuristics

Iterations of Productive Abduction (IAAA) can increase credibility, especially if no other rival explanation is also having that same degree of success. Furthermore, Deduced Productive Abduction (DAAA) can expand the explanatory power of A to additional structures if definite expansions to A's capacities are envisioned.

Iterated Productive Abduction:

If (Only if A's capacities Cs have predicted a series of unexpected W(1-n)) then (an analogous series of unexpected Y(1-n))

Y(1-n)!

If (Only if A's capacities Cs have predicted W(1-n) & Y(1-n)) then (another analogous series of unexpected Z(1-n))

Z(1-n)!

…

So, A(Cs) [IAAA – "Iterated Productive Abduction"]

OR

Deduced Productive Abduction:

If (Only if A's capacities C(n) have predicted a series of unexpected W(1-n)) then (an analogous series of unexpected Y(1-n))

Y(1-n)!

If (Only if A's capacities C(n + 1) have predicted W(1-n) & Y(1-n)) then (another analogous series of unexpected Z(1-n))

Z(1-n)!

…

So, A(Cn +) [DAAA – "Deduced Productive Abduction"]

Both Iterated Productive Abduction and Deduced Productive Abduction can be powerfully credible for scientific explanation. An illustration for the first procedure comes from Maxwell's theory of electromagnetic radiation, which explained the properties of light as manifestations of the same radiating energy found at shorter and longer frequencies (confirmed with radio waves), and explained the properties of both electric and magnetic forces as well, so that a single theory of oscillating electric/magnetic energy obeying a few equations eventually explained a wide range of phenomena. To illustrate the second procedure, consider the concept of the gene, which underwent drastic development during the 20th century. Proposed as the basic unit of heredity passed on the offspring via reproduction, a gene's capacity for transmission and combination with other genes to produce traits in all organisms could additionally explain how a cell's internal processes are regulated if genes also have the ability to control metabolic reactions, suggesting chemical properties for genes. Seeking out those properties in chromosomes, James Watson and Francis Crick ascertained that genes would be stretches of the DNA discovered by X-ray crystallography. Later research has made the concept of 'gene' more complex, as their susceptibility to mutation, reliance on regulatory regions, encoding for multiple proteins, working alongside epigenetic influences, making horizontal transfers (and so on), have explained in succession many puzzling features to cellular activity.

These two procedures are powerful, yet they do go deeper into risky territory. The same problem that emerged with Coduction can arise here for Iterated Productive Abduction: the Explanatory Relationship Problem. We can find analogous features in two distinct matters with enough imaginative creativity. Our efforts might go more into detecting analogous (we imagine) features of structure, and not into the proposed connecting relationships between A, its many capacities, and structures W, Y, Z (etc.). As for Deduced Productive Abduction, matters may be worse because A's capacities are growing during the procedure, so the Explanatory Emptiness Problem can emerge again. If we ignore the issue of maintaining coherent conceptions of the relationships among A and its capacities, this explanatory model may become explanatorily vacuous, and the conception of A eventually seems paradoxical, irredeemably vague, or oddly empty. (The old paradigm of the gene as the unique carrier of information and the powerful initiator of biochemistry is practically extinct, while interest in systems biology and postgenomics grows.[14]) There is no trick to preventing these difficulties in advance. Communities of inquirers must experimentally explore the consequences to expanding an explanation's capacities for the sake of growing its explanatory power, because there is no higher logical method for dictating theoretical modification.

Because there is no higher inferential procedure for dictating modifications to explanations, besides letting them suffer the fate of their own predictive productivity, one way or the other, we are now entirely within the realm of hypothesis experimentation. Each scientific field must rely on the accumulated wisdom of skilled practitioners and useful heuristics for smartly adjusting procedures as inquiries proceed. This is especially the case when a network of interconnected hypotheses form a theory which must undergo further explanatory expansion and testing by risky predictions.

[14] Consult Richardson and Stevens [30].

The familiar problems of deciding which hypotheses within a theory must suffer credibility diminishment or even disconfirmation when things go badly are a matter of scientific heuristics, a higher meta-level problem beyond the scope of strict inferential reasoning.

In the realm of theories—networks of hypothetical explanations about a common matter—the next procedure of Abductive Productive Abduction may be applied, but it is more of an optional heuristic than a required procedure.

AAAA:
Structures W, Y, and Z under experimental conditions EC1!
If (Model A(Cs) can produce W, Y, Z under EC1) then W, Y, and Z
If (Model A(Cs) can produce analogous W, Y, Z under EC2) then analogous W, Y, Z
Analogous W, Y, Z!
…
So, model A(Cs) [AAAA: Abductive Productive Abduction]

With this sketch to AAAA, our stages of procedures must arbitrarily halt to conclude this article. Additional heuristics for modifications to networked models are combinatorially possible. Their patterns can be constructed from the earlier inferential procedures outlines above, by returning to the start of this discussion and letting each instance of 'A' for Abduction stand for Procedural Abduction. Thus, to transform IDA, insert Productive Abduction for that instance of 'A' in 'IDA' to form "inducibly deductive abductively abductive abduction" or just Inducibly Deductive Productive Abduction. A typical scientific field may find a few of these additional heuristics to be practically useful, as it struggles with updating theoretical paradigms and coordinating ontologies with neighboring fields. However useful these advanced heuristics may be, they all still suffer from their characteristic problems and degenerate forms, as warned in previous sections.

8 Conclusion

To summarize, there are five primary categories for sorting the inferential procedures covered by this investigation into the combinatorial possibilities among deduction, induction, and abduction: fallacies, non-scientific procedures, quasi-scientific procedures, scientific procedures, and scientific heuristics. Among the non-scientific and quasi-scientific procedures are found the basic types of mythic thinking and pseudo-scientific thinking, although a discussion about sorting them adequately requires separate treatment.

Fallacies:

A, simple abduction
DA, deducible abduction
IA, inducible abduction

Non-scientific procedures:

AI, abductive induction - Iterative Abduction
AD, abductive deduction - Coduction
DAI, deducibly abductive induction - Retrodicted abduction
ADI, abducibly deductive induction - Retrodicted induction
DIA, deduced inductive abduction - Predicted Abduction
IDA, inducibly deductive abduction - Predictable Coduction
DIA, deducibly inductive abduction - Predictable Abduction
DAI, deducibly abductive induction - Iterative Coduction - religion
IAD, inducibly abductive deduction - Deduced Coduction - theology
AID, abducibly inductive deduction - Reduced Pancosmism - theology

Quasi-scientific procedures:

IDA, inducibly deductive abduction - Predictable Coduction
DIA, deducibly inductive abduction - Predictable Abduction
AIA, abductively inductive abduction - Capacity Overload problem
IAA, induced abductive abduction - degenerates to Deduced Coduction
IIA, induced inductive abduction - Elicited Abduction, Confirmation Biases problem
ADA, abductively deductive abduction - Coductive Abduction, Fraudulent Powers problem

Scientific procedures and heuristics:

DAA, deducibly abductive abduction - Strict Abduction
DDA, deducibly deductive abduction - Deducible Coduction
AAA, abductively abductive abduction - Productive Abduction
IAAA, Iterated Productive Abduction
DAAA, Deduced Productive Abduction
AAAA, Abductive Productive Abduction
(etc.)

References

1. Magnani, L.: Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning. Springer, Dordrecht (2010)
2. Park, W.: Abduction and estimation in animals. Found. Sci. **17**, 321–337 (2012)
3. Musgrave, A.: Deductivism surpassed: or, foxing in its margins? J. Gen. Philos. Sci. **43**, 125–132 (2012)
4. Stanford, P.K.: Damn the consequences: Projective evidence and the heterogeneity of scientific confirmation. Philos. Sci. **78**, 887–899 (2011)
5. Peirce, C.S.: The Collected Papers of Charles Sanders Peirce, vol. 5. In: Hartshorne, C., Weiss, P. (eds.) Harvard University Press, Cambridge, MA (1934)

6. McMullin, E.: The Inference that Makes Science. Marquette University Press, Milwaukee, WI (1992)
7. Niiniluoto, I.: Truth-seeking by abduction. In: Stadler, F. (ed.) Induction and Deduction in the Sciences, pp. 57–82. Kluwer, Dordrecht (2004)
8. Psillos, S.: An explorer upon untrodden ground: Pierce on abduction. In: Woods, J., Gabbay, D., Hartmam S. (eds.) Handbook of the History of Logic, vol. 10, Inductive Logic, pp. 117–151. Elsevier, Amsterdam (2009)
9. Pietarinen, A., Bellucci, F.: New light on Peirce's conceptions of retroduction, deduction, and scientific reasoning. Int. Stud. in the Phil. of Sci. **28**, 353–373 (2014)
10. Campos, D.: On the distinction between Peirce's abduction and Lipton's inference to the best explanation. Synthese **180**, 419–442 (2011)
11. Flach, P., Kakas, A. (eds.): Abduction and Induction: Essays on Their Relation and Integration. Kluwer, Dordrecht (2000)
12. Lipton, P.: Inference to the Best Explanation, 2nd edn. Routledge, London (2004)
13. Paavola, S.: Abduction as a logic of discovery: The importance of strategies. Found. Sci. **9**, 267–283 (2005)
14. Aliseda, A.: Abductive Reasoning: Logical Investigations into Discovery and Explanation. Springer, Dordrecht (2006)
15. Pizzi, C.: Abductive inference and iterated conditionals. In: Magnani, L., Ping, L. (eds.) Model-Based Reasoning in Science, Technology and Medicine, pp. 365–381. Springer, Dordrecht (2007)
16. Schurz, G.: Patterns of abduction. Synthese **164**, 201–234 (2008)
17. Gauderis, T., Van De Putte, F.: Abduction of generalizations. Theoria **75**, 345–363 (2012)
18. Gauderis, T.: Modelling abduction in science by means of a modal adaptive logic. Found. Sci. **18**, 611–624 (2013)
19. Aliseda, A., Beirlaen, M.: A conditional logic for abduction. Synthese **191**, 3733–3758 (2014)
20. Velázquez-Quesada, F.R.: Reasoning processes as epistemic dynamics. Axiomathes **25**, 41–60 (2015)
21. Brigandt, I.: Scientific reasoning is material inference: Combining confirmation, discovery, and explanation. Int. Stud. Philos. Sci. **24**, 31–43 (2010)
22. Lévi-Strauss, C.: Myth and Meaning: Cracking the Code of Culture. Schocken Books, New York (1995)
23. Shackelford, J.: William Harvey and the mechanics of the heart. Oxford University Press, Oxford, UK (2003)
24. Pigliucci, M., Boudry, M.: Philosophy of Pseudoscience: Reconsidering the Demarcation Problem. University of Chicago Press, Chicago (2013)
25. Heidarzadeh, T.: A History of Physical Theories of Comets, From Aristotle to Whipple. Springer, Berlin (2008)
26. Buxton, R. (ed.): From Myth to Reason? Studies in the Development of Greek Thought. Oxford University Press, Oxford, UK (1999)
27. Morgan, K.: Myth and Philosophy from the Presocratics to Plato. Cambridge University Press, Cambridge, UK (2000)
28. Wians, W. (ed.): Logos and Muthos: Philosophical Essays in Greek Literature. State University of New York Press, Albany, NY (2009)
29. Mikalson, J.: Greek Popular Religion in Greek Philosophy. Oxford University Press, Oxford, UK (2010)
30. Richardson, S., Stevens, H. (eds.): Perspectives on Biology after the Genome. Duke University Press, Durham, NC (2015)

Abduction, the Logic of Scientific Creativity, and Scientific Realism

John R. Shook

Abstract A fundamental question for philosophy of science asks, How is knowledge of the world created? A pragmatist approach is constructed to show how discovery and justification are tightly related during the creation of scientific knowledge. Procedural abduction, at the scientific level of Strict Abduction and higher, integrates the learnable (postulations undergoing conceptual development) and the logical (hypotheses undergoing rational scrutiny) quite thoroughly. Discovery and justification are functionally fused together within the organized process of procedural abduction by scientific communities. Four questions posed at the start are answered by this pragmatist philosophy of science as follows. (1) Is scientific creativity methodologically related to scientific justification? Answer: scientific creativity is integral to abductive procedures yielding scientific justification. (2) Can a distinction between genuine science and pseudo-science be clearly defined? Answer: genuine science is distinguished by the application of procedural abduction at the level of Strict Abduction or higher. (3) Does scientific knowledge achieve the legitimacy of scientific realism? Answer: procedural abduction legitimates the credibility of highly-confirmed hypotheses and hence justifies scientific realism. (4) How are scientific communities responsible for establishing scientific knowledge? Answer: scientific communities using procedural abduction realize (in both cognitive and constructive senses) scientific knowledge.

How is knowledge of the world created? Four longstanding issues involved with addressing this general question are usually treated separately by philosophy of science. From a pragmatist approach, there are resolutions to certain issues which additionally yield solutions to the others, and thus all four can be resolved together. These key issues can be expressed as four problematic questions, listed in the order that they are discussed in this chapter.

1. Is scientific creativity methodologically related to scientific justification?
2. Can a distinction between genuine science and pseudo-science be clearly defined?

J. R. Shook (✉)
University at Buffalo, Buffalo, NY 14260, USA
e-mail: jrshook@buffalo.edu

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_10

3. Does scientific knowledge achieve the legitimacy of scientific realism?
4. How are scientific communities responsible for establishing scientific knowledge?

An inadequate answer to one of these problems contributes to making the other problems intractable. Only an examination into the abductive logic of scientific inquiry can show how to resolve all four. The key to these interconnected issues lies with scientific creativity. During the nineteenth and twentieth centuries, dominant theories of scientific methodology ignored creativity or placed creativity's contribution beyond the inferential thinking undergirding scientific credibility. Fixated only on deductive and inductive logic, abductive logic and its creativity has been only thinly considered in epistemology [1] and philosophy of science [2].

Abductive inference has been linked with conjectural creativity in scientific inquiry from its inception with Charles S. Peirce's logical investigations to its elaboration in recent investigations.[1] The pragmatic logic of abductive discovery—with *discovery* bearing the twinned sense of discovering hypotheses having plausibility and discovering that hypothesized entities have reality—requires creativity at every stage from postulation to confirmation. That pervasive role for creativity shows the pragmatist way towards resolving the four issues listed above. Scientific creativity is *integral* to scientific justification; genuine science is distinguished by *procedural* abduction; procedural abduction *legitimates* scientific realism; and scientific communities using procedural abduction *realize* (in both cognitive and constructive senses) scientific knowledge.

1 Types of Scientific Creativity

At the least, scientific creativity is not mere novelty. Like artistic or technical creativity, where innovation develops from earlier forms and designs, the creativity to scientific learning has a largely constructive character. The history of science does not amount to just a loose sequence of novel ideas lacking in cohesion. A plenitude of hypotheses do arise over time, with paradigms twisting and shifting, but a scientific field continually enlarges through discoveries building on discoveries. The culmination of scientific creativity cannot stop short of appreciation and adoption into the growing body of scientific knowledge. Original learning can be surprisingly revelatory but new knowledge must be thoroughly reasonable. Creating knowledge is difficult, and rightly so.

Different types of creativity play important roles in science. On the surface, it is obvious that the knowledge sought by scientific investigators, established as original discoveries, has to be created. Anything empirically known has to be first learned by curious inquirers responsible for learning something new. That learning is created by learners, to supplement and amend knowledge, and then to be subsequently taught as

[1] An entryway into the literature could start by consulting Magnani [18], Paavola [19], Barrena and Nubiola [20], Park [21].

part of that established knowledge. Those instructed learners do not feel so creative; receptivity and flexibility characterize their adoption of knowledge that is new to them. Receptive learning is not a defining component of science itself; creative learning is essential to science. How is scientific knowledge created? A deeper mode of creativity is involved.

Scientific inquiry at minimum enlarges and improves the factual evidence to be considered during an investigation. A body of reliable evidence, no matter how compendious and categorized, needs to be expanded. Investigators can go out to explore and gather fresh material for their study, and they can also return to accumulated evidence for re-inspection and re-interpretation by applying better methods of scrutiny. Both routes exemplify that enlargement of evidence. What seemed evident in the past may later appear less meaningful later, or what seemed uninteresting acquires more significance as overlooked features come into view. Even if already-collected materials are untouched and unchanged, their status as evidence relevant to further inquiry surely changes. New facts are able to arise from old evidence as well as from fresh evidence.

Material evidence, no matter how substantial and abundant, cannot inferentially bear upon justifying any validity to hypotheses—only credible facts could do so. Evidence is "uncovered" as though it pre-exists; interesting facts are surely created. (Pre-existing facts, due to their inadequacy and insufficiency, only provoke those new investigations.) Discovery and creativity are contraries, if their primary meanings are set in direct contrast. What is genuinely discovered cannot also be authentically created—the created thing cannot already exist prior to its creation, while a thing getting discovered must already exist prior to its discovery. And yet, we observe creativity and discovery blending together and intertwining with justification during empirical phases of scientific investigations. Scientific knowledge is created through the process of creating relevant evidence, a process which requires creative engagements with the observable world.

Surprising evidential facts are indispensable, often impelling new thinking and compelling revaluations of older theorizing. Methodical efforts undertaken during the conduct of inquiry create new facts within scientific fields. This view upon creating facts looks contrary to empirical science, which prides itself upon objective methods hostile to human-manufactured "evidence." That much-prized objectivity still involves the creation of new learning, the learning of new facts from enlarging accessible evidence. Furthermore, objectivity implies a reduction of subjectivity, where individual biases flourish. Genuine discovery cannot be merely a fantasy in the mind or a fixation on familiar ground. As an enterprise of discovery, scientific inquiry instead constructs novel conditions where new empirical facts for learning about modeled causes can be openly generated and recorded.[2] This experimental creativity, when accomplished properly, is far from subjective. Reproducibility, repeatability, and robustness across a group of competent investigators are key signs of factual reliability.

[2]Prominent philosophers of science who stress the epistemic link between realistic modelling and controlled experimental conditions include Hacking [22], Geire [23], Cartwright [24].

Creating scientific knowledge relies on creating relevant evidence, which depends on creating experimental conditions that in turn create objective facts—important facts implicated in the creation of credible hypotheses able to creatively accommodate them. During each phase of this discernment of new knowledge through an appreciation of fresh facts and an appraisal of novel hypotheses, a reach of imaginative creativity beyond what is already familiar must be attained. At every level, what has been realistically conceivable, so far, is no longer adequate. Yet, at the same time, whatever is becoming conceivable is also responsible for being reasonable. It is impossible for conception and ratiocination to function in scientific inquiry without continual coordination. That coordination, within procedural abduction, is actually due to their fundamental fusion.

2 Discovery and Justification

How is knowledge of the world created? That creation presumes an integration of what is learnable with what is logical. If learning and logic have nothing in common, not only does their cooperation remain puzzling, but the place for creativity could be divided apart, as if imaginative creativity must stay separated from logical creativity. That manner of subdividing creativity sounds dubious indeed—what gets assigned to "logical creativity" so long as logic is no place for fancy? More commonly, creative discovery gets assigned exclusively to the processes of learning. Intuition, inspiration, imagination—by whatever name, such bursts of creativity seem very different from strict rationality.

As the previous section's tour through primary phases of empirical inquiry has suggested, however, imaginative discovery and inferential justification should be organically unified during the creation of scientific knowledge. If learning and logic are integrated in that common goal of knowledge creation, creativity could not be isolated from reasoning. Each would find its scientific purpose in the other. Creativity would be reasonable, and reasoning would be creative, where a body of scientists are growing a body of knowledge over time.

However timeless the forms of inference may seem, processes of human judgment must be temporal, especially during consideration, consultation, and collaboration. Thinking is temporal through durations; all thoughts have histories. Theories earning their credibility have origins and courses, and even their demises have durable effects in fertilization or fossilization. A scientific body, as a replenishing organization of co-functioning scientists investigating theories over decades and centuries, displays both imaginative creativity and methodic rationality intertwined in intricate harmonies.

A sharp dichotomy between learning and logic establishes a dualism dismembering that organic unity within science. It divides discovery from justification, with spontaneous creativity on one side and strict reasoning on the other side. Creativity would at most have only an external association with logicality, leaving their fruitful relationship as a deep mystery. Why should inferential justification accept intuitive notions as initial inputs for premises, and how would reasoning choose sensible inputs from a plenitude of fancies? Deduction proceeds towards conclusions after initial propositions are granted; it is no business of deduction what ideas get premised. Induction at least demands an array of observed facts before proceeding towards generalizations. Scientific creativity remains a problem where the relationship between learning and logic is a mystery.

Nevertheless, philosophy of science continually distinguishes the context of discovery from the context of justification and then struggles to re-connect them.[3] The post-Kantian separation of empirical contingencies apart from apriori necessities enforced rationalism's dichotomy, and nineteenth century empiricism was no less strict. William Whewell's *The Philosophy of the Inductive Sciences* asserted that the first step beyond the evidence can only be "some happy thought, of which we cannot trace the origin; some fortunate cast of intellect, rising above all rules. No maxims can be given which inevitably lead to discovery. No precepts will elevate a man of ordinary endowments to the level of a man of genius: nor will an inquirer of truly inventive mind need to come to the teacher of inductive philosophy to learn how to exercise the faculties which nature has given him."[4] For Whewell, and so many empiricists claiming expertise over the psychology of knowledge, the insight of a naturally imaginative mind is just an inspirational phase; only logically rigorous inferences can discern true discovery.

Neo-Kantianism and logical empiricism conveyed this view into the early twentieth century, exemplified by Karl Popper. In *The Logic of Scientific Discovery* he stated that "every discovery contains an 'irrational element', or 'a creative intuition'."[5] He expanded on this crucial distinction in this way:

> The initial state, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis not to be susceptible of it. The question how it happens that a new idea occurs to a man – whether it is a musical theme, a dramatic conflict, or a scientific theory – may be of great interest to empirical psychology; but it is irrelevant to the logical analysis of scientific knowledge. This latter is concerned not with questions of fact (Kant's *quid facti?*), but only with questions of justification or validity (Kant's *quid juris?*). Its questions are of the following kind. Can a statement be justified? And if so, how? Is it testable? Is it logically dependent on certain other statements? Or does it perhaps contradict them? … Accordingly I shall distinguish sharply between the process of conceiving a new idea, and the methods and results of examining it logically. As to the task of the logic of knowledge – in contradistinction to the psychology of knowledge – I shall proceed on the assumption that it consists solely in investigating the methods employed in those systematic tests to which every new idea must be subjected if it is to be seriously entertained.[6]

The demise of logical empiricism did not doom Popper's distinction. The larger lesson was amplified: attaining an initial conception is unlike and unrelated to reaching a final conclusion. In general, philosophy must insist that what happens to be *believable* cannot be identified with what should be *credible*. Believability and credibility rest on two separate grounds. Imagining ideas to inspire learning is one

[3]See Nickels [25], Snyder [26], Schickore and Steinle [27].

[4]Whewell [28, vol. 2, p. 186].

[5]Popper [29, p. 32].

[6]Popper [29, pp. 7–8].

process, while justifying learning to count as knowledge is another process. Reducing knowing down to learning violates that distinction and puts psychology in charge of logic (and hence of knowledge and truth too). As Popper well understood, philosophy of science was but one field affected by the broader problem of "psychologism" for philosophical logic [3] and theories of knowledge [4]. Learning is supposed to occur entirely within *natural* psychology while logic is liberated from psychologism by *normative* rationalism.

Narratives about the history of science typically appeal to this creation-justification distinction. One scientist gets credit for first thinking of a new hypothesis, while another scientist is credited with later confirming that hypothesis. Historians of science now understand how scientific advances could not have been so simplistic. The genesis of a hypothesis has receded in significance as theoretical models became more complex, and those abstract models resemble observable things less and less. That oversized role for an individual scientist has also diminished. Behind a complex hypothesis there stands a number of scientists who developed it over time, and teams of scientists are needed for gathering confirmations of that hypothesis. Furthermore, those two processes typically blend and share more in common. The period of development overlaps, and gets involved with, the period of confirmation. Some scientists help to redevelop hypotheses while they participate in designing rounds of experimental trials. A growing body of scientists consult together about the eventual rejection or acceptance of a hypothesis, contributing to the body of knowledge either way.

Allowing how many scientists are typically involved with phases of inquiry, philosophy of science is at least convinced that there is a distinctive logic of justification. In itself, logical justification is not so problematic. Science sets its standards of reasonable inference, to test and justify acceptable hypotheses. On the other hand, the idea of a "logic of discovery" in isolation is harder to conceptualize. Could there be any such thing as a "logic of discovery"?

3 The Learnable and the Logical

The disputed question whether there is a logic of scientific discovery, and wondering how it could relate to the logic of scientific justification, is rooted in the perennial tension between psychology and logic [5]. Modern logic renounced any entanglement with psychology; an understanding of logic requires avoiding the prime fallacy of psychologism. Logic is concerned for the ways that knowledge should be recognized among beliefs. Divorcing the context of discovery from the logic of justification echoes the age-old divide separating learning (temporally psychological) from reasoning (timelessly logical). Actual beliefs and how they happen to form is not supposed to be in logic's department. Knowing, in short, is more than believing. Believed ideas are learned by individuals through the passage of time; known truths are justified by inferences through unchanging norms. How a new idea could inspire original learning must be, it has been claimed, a very different process from the way that an attempt at learning should be justified as knowledge. Intuition, inspiration, imagination—by whatever name, that genesis of creativity by an actual mind seems irreducible to methodical steps for a generic reasoner.

Let logic protest that it truly does guide learning. It is the case that logic must deal indirectly with beliefs, since anything known must at least be believed. For logic, what should be believed is what is learnable, and what is learned should approach knowledge. (What is not knowable, such as the false or mysterious, cannot be learned now, and perhaps never learned.) What is knowable has already been learned, of course—unless something was learned, how could it now be known by anyone?

A. What is knowable must already be learned.

B. What is learned must already be knowable.

If logic has guidance about how knowledge should be learned, how would its guidance be used? It seems as if the known is already the learned, and the learned is already the known—and therefore logic is useless except for its survey of systematized instruction. The object of knowledge is what is already known by some number of minds. This is the basis for deduction: the right conclusion is dependent on reasons, reasons premised and already understood to be acceptable, which reliably guide one's thoughts to the conclusion. The premises must be both familiar and acceptable to one's mind.

C. What is learnable is already conceivable.

Where acceptable premises are to be obtained is not deduction's responsibility. Only premises already accepted as true can yield a knowable conclusion. (Merely hypothetical relations among propositions do not yield known conclusions about something's existence.) The objective of learning is already fully conceived from the start, since a deduction's conclusion is given with the premises. The object of knowledge—indicated by the 'subject' term in the conclusion—is set in the premises, and one's conception of it cannot change while learning from a deduction.

Tenet A can be disputed, for "knowable" has two senses: the knowable is what might become known; or the knowable is already established as known. (When is a river navigable? Only after someone has successfully navigated the length of its waters? Or, is a river navigable before anyone tries? The grammar of '-able' allows both senses.) Potential knowability is distinct from confirmed knowability.

A2. What is potentially knowable may become learned.

B2. What is learned must already be confirmably knowable.

As for logic, it now has a function for learning. What is potentially knowable can become learned through logic's guidance, but that guidance must be cognizant of knowledge's object to some degree. Guidance is no guidance without a conceptualized objective, even if only in vague outline. Furthermore, that guidance must relate this object with information accessible to the learner. Permitting the knowable object to be entirely unlike and unrelated to accessible information is nothing like guidance.

C2. What is potentially knowable is presently conceivable.

If the knowable object is conceived in terms of features evident in accessible information, that conceivability responds to the body of accessible information. If the relevant information changes over time during a period of learning, then the conception of the knowable object can also change. Indeed, intelligent learning modifies conceptions of the knowable object as more and more relevant evidence is gathered. Only unintelligent thought refuses to re-conceive what it is trying to understand.

4 Deduction, Induction, Abduction

If deduction is taken for the paradigmatic mode of inference for knowledge, there can be no logic of discovery. In a sound deduction, the subject of the conclusion—the object to be known—is already accepted as existing when the premises are accepted. Where and why premises are accepted as believable is not deduction's concern. Furthermore, that object of knowledge cannot be changed from the premises to the conclusion; a different 'subject' in the subject-predicate conclusion invalidates a deduction. The pre-given and static status of deduction's object of knowledge explains why deduction yields little learning and no discovery.

Deduction is about learning what is already known, not about the original discovery of something by initial learners. Deduction leads to conclusions of propositional learning about the terms in the premises. This is not empirical learning. Deduction does not conclude anything about the existential discovery of anything. Anything's existence must be presumed in premises. Although a reasoner learns propositions that are new to that learner, only propositions are "discovered." The terms of the conclusion are not new to the reasoner, since the premises must first be understood. Novelty to a term's meaning is unwanted, since a term's meaning should not change between premises and the conclusion. Terms must not change meanings if more premises are added. Through deduction, a term is not discovered, nor is a term's meaning discovered or altered, and nothing that a term may refer to can be discovered. At most, deduction's propositional learning draw attention to relations among understood terms.

Deduction about empirical matters has further restrictions. A learner accepting a conclusion as known accepts the premises as accurate, and accepting an empirical premise involves taking its terms to be about existing matters. Learning an empirical conclusion by deduction is not about discovering a premised term or discovering that a premised thing exists. Nothing in the world is discovered during deductive reasoning.

Induction is, by reputation and results, supposed to be the mode of inference that specializes in original discovery. Learning, if it involves some logicality, requires inferences about (a) objects not already known to exist and (b) not rigidly preconceived. Modes of induction partially satisfy these two criteria for logical learning. Inductive generalizations can anticipate future matters not yet encountered, and they can suggest modified conceptions of things already encountered when conjoined with fresh evidence. For example, the early idea of a microorganism gradually gained specificity as sub-types (such as bacteria, protozoa, and viruses) came into microscopic view, and those classifications themselves developed as more and more organisms were discovered. Induction is restricted by its inability to warrant conceptions of entities impossible to observe by any instrumental means, and limited by its impotence to suggest conceptions of matters quite unlike what has already been observed. Scientific theorizing about non-observable entities, with properties unlike phenomenal qualities, cannot have an entirely inductive basis.

Abductive reasoning is a better model for learning about objects not already known to exist and not familiarly pre-conceived. Abduction introduces and justifies the credibility of fresh hypotheses about unknown things with novel properties, so scientific methodologies require productively abductive theorizing [6] and not just inferences to "the best explanation" [7]. Peirce accordingly claimed that only the original postulations of abduction allows for scientific explanation, with this basic schema:

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

This schema only serves as a comparison with basic forms of deduction and induction. In schematic form, abduction lacks credibility in actual empirical usage, as Peirce himself warned [8]. Abduction in iterative and procedural forms (sketched in following sections) does deliver serious credibility to hypotheses. That credibility can never attain certainty or even confident probability. Valid deduction discerns necessary relations between a conclusion and given premises, while strong induction detects probable conclusions from accumulated premises. Abductive credibility attaches to a surmised conjecture that expects a postulated cause to be responsible for observable effects. Peirce accordingly refers to "deductive necessity," "inductive probability," and "abductive expectability" ([4], 5.194).

The creativity inherent to abductive postulation, as Peirce repeatedly explained, allows for a genuine logic of discovery [9]. In this logic for learning, that static "Discovery-Justification" dichotomy separating learning from logic is replaced by a functional "Postulation-Confirmation" distinction within a unified process of reasoned discovery.

5 Abduction and Postulation

In 1878, Peirce published the sixth part of his "Illustrations of the Logic of Science" titled "Deduction, Induction, and Hypothesis." By "hypothesis" Peirce was referring to what he also called "retroduction" and later labeled as abduction [10]. On deduction, Peirce points out that it "adds nothing to the premises, but only out of the various facts represented in the premises selects one and brings the attention down to it" ([4], 2.643). Comparing induction with hypothesis (abduction), he writes,

> By induction, we conclude that facts, similar to the observed facts, are true in cases not examined. By hypothesis, we conclude the existence of a fact quite different from anything observed, from which, according to known laws, something observed would necessarily result. The former, is reasoning from particulars to the general law; the latter, from effect to cause. ([4], 2.536)

Induction can ascertain patterns and regularities among things sharing similarities. Discovering a not-yet-observed explanation responsible for those matters asks creative thinking to go beyond induction.

As Peirce refined and enlarged his approach to abduction, he continually emphasized science's essential dependence on abduction's creativity, transcending any observational reach.

> All the ideas of science come to it by way of abduction. Abduction consists in studying facts and devising a theory to explain them. ([4], 5.145)

> Abduction is the process of forming explanatory hypotheses. It is the only logical operation which introduces any new idea. ([4], 5.172)

The relationship of abduction's creativity with confirmation is left unclear by these brief statements.

The simplest formulation of abduction is, as Peirce well knew, just a formal fallacy of affirming the consequent. That concise schema only mentions the postulated entity once in the two premises: "But if A were true, C would be a matter of course." Where does that conception of A come from? It does not arrive from somewhere beyond abduction since it is a component of abductive reasoning. Yet its singular mention in the premises makes it look like it descends from clouds of imagination.

Peirce does say that "the abductive suggestion comes to us like a flash. It is an act of insight" ([4] 5.181). However, Peirce treats this "insight" more like an informed guess [11] that arises in various guises *during* inferential reasoning [12]. He wrote, "It must be remembered that abduction, although it is very little hampered by logical rules, nevertheless is a logical inference asserting its conclusion only problematically or conjecturally, it is true, but nevertheless having a perfectly definite logical form" ([4] 5:188). There is no contradiction between these two statements about abduction, unless one (wrongly) presumes that an initial insight is never modified throughout the process of abductive reasoning towards its eventual conclusion. That presumption is essential to valid deduction (avoiding the fallacy of four terms), but Peirce did not reduce abduction to a sort of deductive argument.

Abduction in the hands of scientific inquiry is never just simple abduction in pure form. Induction and abduction (hypothesis) cooperate in concert, according to Peirce.

> The great difference between induction and hypothesis is, that the former infers the existence of phenomena such as we have observed in cases which are similar, while hypothesis supposes something of a different kind from what we have directly observed, and frequently something which it would be impossible for us to observe directly. Accordingly, when we stretch an induction quite beyond the limits of our observation, the inference partakes of the nature of

> hypothesis. It would be absurd to say that we have no inductive warrant for a generalization extending a little beyond the limits of experience, and there is no line to be drawn beyond which we cannot push our inference; only it becomes weaker the further it is pushed. Yet, if an induction be pushed very far, we cannot give it much credence unless we find that such an extension explains some fact which we can and do observe. Here, then, we have a kind of mixture of induction and hypothesis supporting one another; and of this kind are most of the theories of physics. ([4] 2.640.)

Furthermore, the explanatory power of abduction also includes deduction:

> Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea; for induction does nothing but determine a value and deduction merely evolves the necessary consequences of a pure hypothesis. Deduction proves that something *must* be, Induction shows that something *actually is* operative, Abduction merely suggests that something *may be*. Its only justification is that from its suggestion deduction can draw a prediction which can be tested by induction and that, if we are ever to learn anything or to understand phenomena at all, it must be by abduction that this is to be brought about. ([4] 5:171)

Reaching a conclusion earning abductive credibility is the result of prolonged inquiry incorporating phases of creative postulation together with induction and deduction.

A closer examination of abductive procedures for science, first elaborated in Shook [13] and sketched in the next section, reveals how they require a dynamic relationship between the accumulation of new empirical evidence and the alterations needed to the conception of the object of knowledge proposed in a hypothesis. That dynamic relationship between evidence and hypothesis accounts for the scientific realism that arises from abduction. Successful confirmations from abductive procedures yield conclusions credibly affirming the real existence of their hypothesized objects of knowledge.

6 Abduction and Confirmation

Deduction, induction, and abduction can be simplistically formulated in their pure timeless forms. Imitating Peirce's examples, consider the fruit of a particular tree.

Deduction – Atemporal

Fruits from that tree are red.

These fruits are from that tree.

Therefore, these fruits are (surely) red.

Induction - Atemporal

These fruits are from that tree.

These fruits are red.

Therefore, fruits from that tree are (probably) red.

Abduction - Atemporal

That tree's fruit is red.

If these fruits are from that tree, then they are red.

Therefore, these red fruits are (possibly) from that tree.

Their atemporal forms allow for schematic comparison, to show how none of them are reducible to another form.

Treating abduction only as a straightforward sort of premise-to-conclusion reasoning with but two premises is misleading. During the procedures of complex types of abduction, the object of the conclusion is re-conceived during the consideration and re-consideration of additional sought-for premises. The point of abductive reasoning is to improve conceptions of that postulated object in its capacity to be causally responsible for observed effects, while the plausibility of its efficacious reality grows in relation to an enlarging evidence base. The eventually discovered object is not already fully conceived from the start.

Peirce expected that the three kinds of inference—deduction, induction, and abduction, should cooperate in empirical discovery. His 1903 Harvard *Lectures on Pragmatism* says:

> Abduction merely suggests that something may be. Its only justification is that from its suggestion deduction can draw a prediction which can be tested by induction, and that, if we are ever to learn anything or to understand phenomena at all, it must be by abduction that this is to be brought about. ([4] 5.17)

Peirce occasionally referred to "mixed" reasonings and inferences ([4] 2.774, 2.787, 7.218). He emphasized how deduction, induction, and abduction are distinct components in science, unable to perform another's inferential work.

> Nothing has so much contributed to present chaotic or erroneous ideas of the logic of science as failure to distinguish the essentially different characters of different elements of scientific reasoning; and one of the worst of these confusions, as well as one of the commonest, consists in regarding abduction and induction taken together (often mixed also with deduction) as a simple argument. Abduction and induction have, to be sure, this common feature, that both lead to the acceptance of a hypothesis because observed facts are such as would necessarily or probably result as consequences of that hypothesis. But for all that, they are the opposite poles of reason, the one the most ineffective, the other the most effective of arguments. The method of either is the very reverse of the other's. Abduction makes its start from the facts, without, at the outset, having any particular theory in view, though it is motived by the feeling that a theory is needed to explain the surprising facts. Induction makes its start from a hypothesis which seems to recommend itself, without at the outset having any particular facts in view, though it feels the need of facts to support the theory. Abduction seeks a theory. Induction seeks for facts. In abduction the consideration of the facts suggests the hypothesis. In induction the study of the hypothesis suggests the experiments which bring to light the very facts to which the hypothesis had pointed. ([4] 7.218)

Deduction, induction, and abduction have very different inferential characters and results. That is why each needs the others for productive and predictive inquiries. For example, Peirce recounts how abduction and induction can cooperate during investigations into explanations for empirical patterns:

> Presumption, or, more precisely, abduction … furnishes the reasoner with the problematic theory which induction verifies. Upon finding himself confronted with a phenomenon unlike

> what he would have expected under the circumstances, he looks over its features and notices some remarkable character or relation among them, which he at once recognizes as being characteristic of some conception with which his mind is already stored, so that a theory is suggested which would explain (that is, render necessary) that which is surprising in the phenomena. ([4] 2.776)

In a 1902 manuscript, after declaring that "arguments are either deductions, inductions, abductions, or mixed arguments," Peirce describes a thoughtful process that mixes abduction with induction.

> Suppose, then, that, being seated in a street car, I remark a man opposite to me whose appearance and behavior unite characters which I am surprised to find together in the same person. I ask myself, How can this be? Suppose I find this problematic reply: Perhaps he is an ex-priest. He is the very image of such a person; he presents an icon of an ex-priest. Here is an iconic argument, or abduction of it. Secondly, it now occurs to me that if he is an ex-priest, he should be tonsured; and in order to test this, I say something to him calculated to make him take off his hat. He does so, and I find that he is indeed tonsured. Here at last is an indication that my theory is correct. I can now say that he is presumably an ex-priest, although it would be inaccurate to say that there is any definite probability that he is so, since I do not know how often I might find a man tonsured who was not an ex-priest, though evidently far oftener than he would be one. The supposition is, however, now supported by an inductive induction, a weak form of symptomatic or indexical argument. It stands on a widely different basis from that on which it stood before my little experiment. Before, it rested on the flimsy support of similarity, or agreement in "flavor." Now, facts have been constrained to yield confirmation to it by bearing out a prediction based upon it. Belief in the theory rests now on factual reaction to the theory. [14]

Peirce's story illustrates an inquiry that generates new evidence from an abductive guess which in turn supports the plausibility of that hypothesis. Confirming evidence is not independent from the postulated hypothesis; that evidence may never have been sought and found without such a hypothesis in mind. The notion of a hypothesis generating its own evidence must look suspicious from the standpoint of static deductive logic or sequenced inductive logic. Abductive reasoning is circular, in the sense that the growing quality of the evidence is the responsibility of the hypothesis's greater explanatory power. Atemporal reasoning schemas cannot reproduce or license such a mutually supportive relationship between postulation and confirmation stages.

Learning takes time; learning through reasoning is assuredly temporal. Imagining, thinking, and predicting are mental processes having durations. Peirce typically depicts induction and abduction as thoughtful procedures extended over time. Basic forms of inductive and abductive procedures can accordingly be schematized.

Induction - Temporal

These 3 small fruits are from that tree.

    Those fruits are also red.

These 4 small red fruits are from that tree too.

    Those fruits are also sweet.

These 5 small red sweet fruits are from that tree too.

Those fruits are also soft.

Therefore, fruits of that tree are small, red, sweet, and soft.

During the process of temporal induction, one's conception of the conclusion's object, that tree's fruits, is modified. Alterations to the object of the conclusion also occur for temporal abduction.

Abduction - Temporal

These are small and red fruits.

That tree's fruit is small and red.

If these fruits are from that tree, then they are small and red.

These small red fruits are also sweet.

That tree's fruits are also sweet.

Therefore, these fruits are from that tree.

And, as Peirce proposes, abduction and induction in their temporal forms can be combined.

Abduction – Inductively Temporal

These are red fruits.

That tree has red fruit.

If those fruits came from that tree, then they would be red.

These same red fruits are also small.

That tree has small red fruit.

If those fruits came from that tree, then they would be small and red.

These same small red fruits are also sweet.

That tree has small red sweet fruit.

If those fruits came from that tree, then they would be small, red, and sweet.

Therefore, these small red sweet fruits came from that tree.

Inductively temporal abduction ensures that one's conception of that tree's fruits is gradually modified, and the actual origin of those fruits (that tree) is now expected by the reasoner. Furthermore, with each additional observation, confidence in the accuracy of this conclusion reasonably increases. Deduction is not left out of this iterative process. At each stage, the statement of the hypothesis is a deduction in miniature, e.g.: "If those fruits came from that tree, then they would be small and red." Gathering more empirical evidence modifies the conception of the conclusion's object, and it develops the hypothesis. Let us call this "procedural abduction." The dynamic relationship between the growing evidence and the developing hypothesis is the basis for the realism that arises from procedural abduction: it is more and more credible that the hypothesized entity exists. This is discovery, from a reasoning procedure, where no bright line is separating the logic apart from the learning.

In summary so far, the inferential modes of deduction, induction, and abduction can be compared in stages from postulation to confirmation.

Deduction does not seek more premises, deduction cannot change the meaning of terms during reasoning, and deduction cannot discover the existence of anything.

Induction could seek more premises, and repetitive induction can change the meaning of terms during reasoning, but induction does not discover things with novel properties.

Abduction should seek out more premises, abduction must change the meaning of terms during reasoning, and abduction can discover unfamiliar things with novel properties. Furthermore, iterative abduction can raise the level of reasonable confidence in the real existence of those things.

Although Peirce labelled a proposal of a hypothesis as an "abduction" it would be a mistake to isolate scientific creativity in general within abduction alone, apart from deduction and induction. Peirce never made that mistake. Only the combination and integration of the three forms of inference gets productively utilized within empirical inquiry.

7 Procedural Abduction

Peirce offered a few examples of cooperation among forms of inference, but he did not explore mixed inferences further. Many combinations of deduction, induction, and abduction can be formulated, and some of them inform sound scientific methodologies. Twenty-five combinations are delineated in Shook [13], ranging from the fallacious and pseudo-scientific to the proto-scientific and fully scientific. Four types of reasoning, from simpler to more complex forms, serve to illustrate here how the last type of scientific abduction, "strict abduction," is able to warrant credible conclusions about postulated entities. Qs, Rs, and Ss are placeholders for any sort of observed phenomena, while A (and its capacities C1, C2, etc. that make a difference to observable evidence) is a placeholder for any postulated entity (e.g. an object, model, energy, force, field, and so on).

Retrodicted Induction

Qs!

Suppose that If A then Qs [now expecting Qs from A's vague definition]

Rs!

Suppose that If A then Rs [now expecting Rs from A's vague definition too]

Ss!

Suppose that If A then Ss [now expecting Ss from A's vague definition too]

…

So, A

Retrodicted Induction superficially looks like an abductive procedure. It is far more suspicious, because A's definition is designed in advance to 'explain' not just some initial Qs but also plenty of other vaguely indicated matters, so that any chosen Rs, Ss, and Ts (etc.) can get 'explained' when they show up later. Retrodicted Induction cannot attain the level of scientific theorizing.

Predicted Abduction

If A then Qs [from a vague idea of A, by deduction Qs would be expected from A]

A pattern of Qs gets discovered!

If A then Rs [from a vague idea of A, by deduction Rs would be expected from A]

A pattern of Rs gets discovered!

…

So, A

Predicted abduction also falls short of the level required for fully scientific theorizing. It allows a thinker to remain stubbornly attached to an initial conception of the entity to be discovered.

Predictable Coduction

If A then Qs have features F1 [given A's definition, by deducing how Qs having F1 are expected]

Qs have F1!

If A then Rs also have analogous features F2 [after adjusting A's definition, then deducing how Rs having F2 are expected, while still deducing Qs with F1 too]

Rs have F2!

…

So, A

Predictable Coduction is more plausible, because A is well-defined rather than vague, and A's definition is permitted to developed in only incremental ways in response to evidence. Predictable Coduction lacks explanatory plausibility, however, since it actually only "explains" things as they get discovered.

Strict Abduction

Qs!

Suppose (only if A has C1), then Qs

Suppose (only if A has C1-2), then Qs & Rs

Rs!

Suppose (only if A has C1-3), then Qs, Rs & Ss

Ss!

…

So, A(Cn)

At this level of scientific theorizing, where postulation and confirmation are thoroughly intertwined, it is no longer an easy matter to see where logic and learning are divided apart. The conceptual creativity applied to developing the object of the conclusion is not a separate thought process apart from the inferential rationality that eventually warrants acceptance of that entity's existence.

8 Abductive Scientific Realism

The question of philosophical realism is a metaphysical issue, unlike scientific realism. Even if one grants a measure of scientific realism, affirming that postulated entities with ample scientific confirmation are credibly real (more or less as theories conceive them), philosophy can still ask its skeptical question, "Is it rational to think that science's affirmed entities actually exist?" Science's most confirmed entities may not be actually knowable, if philosophy knows knowledge better than any amount of science. Metaphysical anti-realism can be compatible with modest scientific realism, if only to warn science that its excusable confidence in theoretical entities cannot determine their actual reality or compel a rational mind to take them as truly real. This philosophical anti-realism sets the bar for knowledge higher than any methodological standards followed during scientific inquiry. Philosophical naturalism, by contrast, takes the position that science's highly confirmed entities should enjoy at least as much credibility (and often more) as anything else familiarly known from experience [15].

This chapter's topic is scientific realism, not metaphysical realism/anti-realism, or philosophical naturalism. The motivating question is, "Is it reasonable for scientific realism to be affirmed in the course of empirical inquiries applying sound scientific methodologies?" When the application of procedural abduction in scientific methodologies is considered, then scientific realism is warranted because scientific hypotheses about postulated entities become credibly reasonable.[7] Standing outside of science, and pondering how to inductively or abductively justify scientific realism, is already philosophically futile and scientifically irrelevant [16]. The best explanation for science's success is science's own work: if science itself does not sufficiently justify the credibility of its confirmed hypotheses in the first place, nothing can. Fortunately, science has no need of any non-scientific or metaphysical assistance. Naturalism's worldview, for example, is plausible only if scientific realism is already reasonable; nothing about the scientific realism due to procedural abduction needs any axiom or premise of naturalism.

The credible plausibility to abductive scientific realism lies in the special features of Strict Abduction and higher-order abductive procedures in Shook [13]. Crucial features have to do with the creative postulation and re-conceptions of the entity to be discovered. Loose ideas of entities allow for vague predictions about amorphous evidence, evidence that any number of similarly imprecise postulations could equally well "explain". The poor reputation of abduction is not due to abductive reasoning itself, but rather to vague and unrevised ideas of postulated causes. Strict abduction deals with a postulated entity A by exercising tight control over modestly modifying the conception of A during the reasoning process. That control is emphasized in the first two special features of Strict Abduction.

> First, at each stage, the conception of A has only one clear definition and set of capacities. Only the capacities required to account for the phenomena are attributed to A, and whatever

[7] Recent studies linking scientific realism with abductive inquiry include Magnani and Betolotti [30], Niiniluoto [31].

the definition of A may be, that definition is only permitted to be compatible with those Cs applied in the procedure.

Second, no other conceptions of A, beyond those Cs proposed to account for Qs, Rs, Ss (and so on) are regarded as relevant. Strict Abduction does not permit the definition of A to range beyond whatever is minimally necessary for it to have its explanatory capacities. That strict control allows successful predictions to more impressively support the postulated hypothesis.

Any responsibility for the vagueness or precision of conception of postulated entities must rest with the human conceivers, not the entities. The fault lies with scientists for failing to better define and refine their hypotheses, thereby permitting undeserved "confirmations" and allowing unscientific theories to proliferate. It is a mistake to depict scientific inquiry as a thought process undertaken by a solitary thinker. Peirce expected a scientific community to conduct and control the scientific enterprises of empirical inquiry and collectively evaluate their results. Scientific communities yield knowable discoveries, not any lone mind.

9 Scientific Communities

Procedural abduction works best for a community of scientific inquirers who consult together about how realistic a hypothesis can become, while they enlarge the collection of evidence and simultaneously develop their conceptions of postulated entities. These additional features of procedural abduction, exclusively the responsibility of scientific communities, have essential roles:

Third, due to the bounded clarity supplied by the second feature, a community of inquirers can apply A together and everyone can agree upon what the explanation is and what it so far entails.

Fourth, although a community will have disagreements over what new capacities A should have for increasing its predictive range, both the current definition of A and the presently assigned capacities place compatibility constraints on the sort of new capacities that can be assigned to A.

Fifth, if a new prediction goes badly, the community of inquirers only needs to doubt the new implicated capacity of A, not the rest of the capacities of A, preserving the explanatory power A had already earned.

Sixth, the expansion of A's capacities and its explanatory range can halt and pause whenever the community finds no work for A to do presently, but A can be put to work again in the future when opportunities come for relevant observations.

More complex kinds of procedural abduction than Strict Abduction all share in these six features. Those features prevent a hypothesis from being able to explain far too much, and from trying to explain new phenomena only after they are observable. All the same, a hypothesis explaining too much too easily can seem convincingly realistic to the smartest minds, including scientists. Histories of scientific fields are replete with tales about good scientists who stubbornly cling to their inadequate hypotheses. Humility is perhaps the prime virtue of scientific character. (Peirce pointed to scientific analogues of faith, hope, and charity as well; see Shook [17]). Peirce wrote,

> The scientific world is like a colony of insects in that the individual strives to produce that which he himself cannot hope to enjoy. One generation collects premises in order that a distant generation may discover what they mean. ([4]. 7.87)

Since scientific knowledge of the real world is created, something in this world accomplishes that knowledge—the community of scientific inquirers, who have a shared history of discovery and a shared future of hypothesis testing, bound together by a commitment in their common purpose of creating knowledge. Peirce explicitly connected the ideal of the scientifically real with the idea of the scientific community.

> The real, then, is that which, sooner or later, information and reasoning would finally result in, and which is therefore independent of the vagaries of me and you. Thus, the very origin of the conception of reality shows that this conception essentially involves the notion of a COMMUNITY, without definite limits, and capable of a definite increase of knowledge. ([4] 5.311)

That growth of discovered knowledge is due to abductive procedures applied by scientific inquirers. Procedural abduction, by maximizing the value of evidential information and inferential reasoning, yields discovery in its genuine sense of scientific realism.

Procedural abduction overcomes that long-standing dichotomy between psychological learning and rational logic. Where learning and logic, and discovery and justification, are unified for the inclusive goal of knowledge creation, creativity could not be isolated from reasoning. Each finds its scientific purpose in the other. Creativity is reasonable, and reasoning is creative, where an organization of scientists are growing organized knowledge. Three modes of creativity have come up in this scientific context: novelty, development, and organization.

> Novelty – new things one after another after another. However, mere novelties may not be relevant to each other, so development is needed.
>
> Development – enlarging capacities to effectively manage sequenced novelties. However, independent developments are not automatically coordinated with each other, so organization is needed.
>
> Organization – improving integration of the whole through harmonious co-development. However, only committed organizations with a shared history and future can guarantee this co-development, so scientific community is needed.

This chapter asked a fundamental question for philosophy of science: How is knowledge of the world created? It was proposed that what is learnable and what is logical is integrated and unified by the processes of creating knowledge. This would require that discovery and justification are organically unified during the creation of knowledge. Procedural abduction, at the scientific level of Strict Abduction and higher, integrates the learnable (postulations undergoing conceptual development) and the logical (hypotheses undergoing rational scrutiny) quite thoroughly. This is where discovery and justification are functionally fused together within the organized process of procedural abduction by scientific communities.

The four questions posed at the beginning are answered by this pragmatist philosophy of science as follows. (1) Is scientific creativity methodologically related to scientific justification? Answer: scientific creativity is integral to abductive procedures yielding scientific justification. (2) Can a distinction between genuine science and pseudo-science be clearly defined? Answer: genuine science is distinguished by the application of procedural abduction at the level of Strict Abduction or higher. (3) Does scientific knowledge achieve the legitimacy of scientific realism? Answer: procedural abduction legitimates the credibility of highly-confirmed hypotheses and hence justifies scientific realism. (4) How are scientific communities responsible for establishing scientific knowledge? Answer: scientific communities using procedural abduction realize (in both cognitive and constructive senses) scientific knowledge.

References

1. Hintikka, J.: What is abduction? The fundamental problem of contemporary epistemology. Trans. Charles S. Peirce Soc. **34**, 503–533 (1998)
2. Aliseda, A.: The logic of abduction: An introduction. In: Magnani, L., Bertolotti, T. (eds.) Handbook of Model-based Science, pp. 219–230. Springer, Cham, Switzerland (2017)
3. Brockhaus, R.: Realism and psychologism in 19th century logic. Philos. Phenomenological Res. **51**, 493–524 (1991)
4. Peirce, C. S.: The collected papers of Charles Sanders Peirce, vol. 8. In: Hartshorne, C., Weiss, P, Burks, A.W. (eds.) Harvard University Press, Cambridge, MA (1931–1935, 1958). (Citations are to volume and paragraph numbers)
5. Peckhaus, V.: Psychologism and the distinction between discovery and justification. In: Schickore, J., Steinle, F. (eds.) Revisiting Discovery and Justification, pp. 99–116. Springer, Berlin (2006)
6. Pietarinen, A., Belluci, F.: New light on Peirce's conceptions of retroduction, deduction, and scientific reasoning. Int. Stud. Philos. Sci. **28**, 353–373 (2014)
7. Yu, S., Zenker, F.: Peirce knew why abduction isn't IBE: a scheme and critical questions for abductive argument. Argumentation **32**, 569–587 (2018)
8. Plutynski, A.: Four problems of abduction: a brief history. HOPOS: J. Int. Soc. Hist. Philos. Sci. **1**, 227–248 (2011)
9. Anderson, D.R.: Creativity and the Philosophy of C. S. Peirce. Martinus Nijhoff, Dordrecht (1987)
10. Anderson, D.R.: The evolution of Peirce's concept of abduction. Trans. Charles S. Peirce Soc. **22**, 145–164 (1986)
11. Tschaepe, M.: Guessing and abduction. Trans. Charles S. Peirce Soc. **50**, 115–138 (2014)
12. Tschaepe, M.: Gradations of guessing: preliminary sketches and suggestions. Contemp. Pragmatism **10**, 135–154 (2013)
13. Shook, J.R.: Abduction, complex inferences, and emergent heuristics of scientific inquiry. Axiomathes **26**, 157–186 (2016). Reprinted in this volume
14. Peirce, C. S.: The Carnegie application for logic research, Draft E" MS L75.163-173, held by The Peirce Edition Project at Indiana University Indianapolis (1902). Accessed online at http://arisbe.sitehost.iu.edu/menu/library/bycsp/L75/ver1/l75v1-06.htm
15. Shook, J.R.: Varieties of twentieth century American naturalism. Pluralist **6**, 1–17 (2011)
16. Worrall, J.: The no miracles intuition and the no miracles argument. In: Dieks, D. (ed.) Explanation, Prediction, and Confirmation, pp. 11–21. Springer, Dordrecht (2011)
17. Shook, J.R.: Reasoning is communal in method and spirit. In: Thellefsen, T., Sorensen, B. (eds.) Charles Sanders Peirce in His Own Words, pp. 70–83. Mouton De Gruyter, Berlin (2014)

18. Magnani, L.: The Abductive Structure of Scientific Creativity. Springer, New York (2017)
19. Paavola, S.: On the Origin of Ideas: An Abductivist Approach to Discovery, 2nd edn. LAP Lambert Academic Publishing, Saarbrücken, Germany (2012)
20. Barrena S., Nubiola J.: Abduction, the logic of creativity. In: Bloomsbury Companion to Contemporary Peircean Semiotics, pp. 185–203. Bloomsbury, London (2019)
21. Park, W.: Abduction in Context: The Conjectural Dynamics of Scientific Reasoning. Springer, Cham, Switzerland (2017)
22. Hacking, I.: Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. Cambridge University Press, Cambridge, UK (1983)
23. Giere, R.N.: Science Without Laws. University of Chicago Press, Chicago (1999)
24. Cartwright, N.: Nature's Capacities and Their Measurement. Oxford University Press, Oxford, UK (1989)
25. Nickles, T.: The strange story of scientific method. In: Meheus, J., Nickles, T. (eds.) Models of Discovery and Creativity, pp. 167–207. Springer, Dordrecht, Germany (2009)
26. Snyder, L.J.: Hypotheses in 19th century British philosophy of science: Herschel, Whewell, Mill. In: Heidelberger, M., Schiemann, G. (eds.) The Significance of the Hypothetical in Natural Science, pp. 59–76. Walter de Gruyter, Berlin (2009)
27. Schickore, J., Steinle, F.: Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction. Springer, Berlin (2006)
28. Whewell, W.: The Philosophy of the Inductive Sciences. J.W. Parker, London (1840)
29. Popper, K.: The Logic of Scientific Discovery. Routledge, London (1959)
30. Magnani, L., Bertolotti, T. (eds.): Springer Handbook of Model-based Science. Springer, Cham, Switzerland (2017)
31. Niiniluoto, I.: Truth-Seeking by Abduction. Springer, Cham, Switzerland (2018)

Human Reasoning and Theories of Cognition

Abduction and the Logic of Inquiry: Modern Epistemic Vectors

Jay Schulkin

Abstract CS Peirce introduced the concept of abduction into our epistemic lexicon. It is a view of problem solving that emphasizes ecological contexts, preparatory or predictive predilection knotted to learning and inquiry. Abduction is essentially tied more broadly to pragmatism. One view of the brain reflects the fact that predictive predilections knotted to abduction or hypothesis testing dominates the landscape of diverse forms of problem solving. Abduction is biologically constrained and contextual, not a monolithic term and runs the range of neural capability.

Keywords Abduction · Inquiry · Disruption · Inference · Cephalic · Conservation

1 Introduction

Problem solving is at the root of foraging for epistemic coherence; its roots in its diverse forms is biological. We come ready to investigate, to search for coherence and meaning. Embedded in the process of satisfying our basic wants and needs are a number of problem-solving skills—something I like to call "cephalic capabilities." I use the word cephalic in neuroscience [1–3], an older word not much used in the neurosciences, because of its link to the brain in the context of the body. The living brain is not detached from action and divorced from the contours of adaptation; rather, a brain is part of a body in space and time that is the rare limiting factor for behavior and experience.

Moreover, we do not come into the world as a blank slate. Abduction is in a context of prior adaptive resolutions, prior probabilities in predictive machinery inherent in the design of the nervous system, diverse forms of cognitive capabilities permeate neuronal expression and structure at the very onset of life [4]. These capabilities include the ability to think about numbers, food, space, time, detecting others and their beliefs, etc. We also start out with a certain cephalic plasticity.

J. Schulkin (✉)
Department of Neuroscience, Georgetown University, Washington, USA
e-mail: jjs54@georgetown.edu

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_11

Cephalic systems traverse the whole of sensory and information-gathering systems with which we sample the world and update our orientations. There is no separation of a mind from a body, but diverse forms of information sampling of the environment, both internal for what is to be sustained in viable ways, and sampling external events for what is occurring [2]. Indeed, core abductive moments "getting the gist" of something fast [5–7] makes sense from a neuroscience perspective by the massive neural connectivity between neocortical or amygdala innervation of sites in the brainstem or the peripheral visceral nervous systems that regulate the alimentary systems, namely the gut those hunches, generating an idea, those abductive moments, are rich in gut//visceral/cortical neural connectivity.

Bodily sensibility and the issue of cephalic functioning (and its evolution by adding on diverse forms of information processing, of sampling and appraising diverse terrains) is its adaptive value and applications [8, 9] and continued expression that underlies our epistemic engines of discovery [10, 11]. In other words, problem solving is a continuous function across the natural and cultural landscape as we look into our evolutionary past, neural systems, and present circumstances [12], it is mostly local and rooted in context and need.

In this chapter, I start with giving a sense of what abduction is about, and then some of the resources that figure in active inference, abduction and problem solving.

Abduction and Inquiry: CS Peirce suggested that there are no precognitive events but rather degrees of cognitive action underlying perception, attention, and action that runs through our experience. Indeed, we know that human organization is based largely on anticipatory cognitive system tied to ecological contexts of relevance and neural design.

For Peirce, pragmatism is knotted inextricably to *abduction*. By abduction, Peirce meant the genesis of ideas or, in other words, hypothesis formation and testing. Abductive events are at the heart of inquiry. For Peirce, abduction is "preparatory;" and induction is a "concluding step" grounded in natural outcomes and deduction. By, introducing the idea of abduction, Peirce offered an alternative to rationalism and deduction, as well as endless empiricism and induction.

Anticipation, a response to the awareness of time, is a fundamental feature of our cognitive capabilities; it is tied to everything about us. The endogenous clocks are reminders in preparation for events, in anticipation of action. They are the origins of forethought [13], abductive inferences begin to emerge with the anticipation of events and their breakdown and within a context of cephalic/ecological and social resources.

Our evolutionary orientation toward familiar objects remains somewhat constant but our understanding always reflects our specific culture as well. We come particularly prepared to focus on animate objects, like body parts, including that of the brain [14].

We come prepared to forage for the coherence of our neural processes as much as we come prepared to look for food. Our evolutionary success is contingent upon both. Knowledge acquisition is constrained by context and expectations; the law of "sufficient reason," for instance, is contextualized by investigators, the tools for inquiry that are available, and the theories that dominate the inquiry.

A core capacity for investigations is rooted in this basic need to find coherence, to sample for stable, viable, predictable events, from which to build action-oriented anticipated functions and foraging for coherence in inductive machinery. Multiple diverse exploratory processes are embedded in this cognitive/neural capability.

Indeed, Peirce was insightful. Peirce understood that “thinking is a species of the brain and cerebration is a species of the nervous action” (Peirce 1880/1992, Vol. 1 p. 107). Peirce always noted that there are no precognitive events but rather degrees of cognitive action underlying perception, attention, and action [6]. Indeed, we know that human organization is based largely on anticipatory cognitive systems, most of which encompass the vast cognitive unconscious [15]. When Peirce correctly noted, “It was impossible to know intuitively that a given cognition is not determined by a previous one” [16, p. 30], he was close to recognizing that cognitive resources figure in the organization of action, as do the larger habits and cultural expression, something some 30 years later John Dewey would note in his critique on the “reflex arc” [8, 17, 18].

Peirce regarded pragmatism as the method to determine larger consequences. This represented a move away from nominalism towards the realism of Scotus and Reid about kinds of objects and our knowledge of them. For Peirce, pragmatism is knotted inextricably to *abduction* [7]. By abduction, Peirce meant the genesis of ideas or, in other words, hypothesis formation and testing. Abductive events are at the heart of inquiry (Table 1).

Abduction is “preparatory;” and induction is a “concluding step” grounded in natural outcomes [7, 19, 20]. Hypothesis formation as it related to the science of discovery [21, 22, 23]. There was no detached starting point for the eradication of all doubt as there was in Descartes (see also see [24]). Peirce underlie epistemic activities and the cumulative results of inquiry. He added the additional element of something akin to an instinct hitting on the right idea: abduction. Here evolutionary success is an important element and the larger cultural expression that cultivates successful problem solving and discovery.

Hypothesis formation is essential for drawing inferences in the context of deducing consequences and inducing noted relationships. The causes and the consequences are embedded in larger based frameworks [22, 25]. Abduction offered entrenchment in the knowing process tied to prediction and coherence. And predictive coherence are some of the features that the living brain depends on: reliabilism [11, 26–28].

Table 1 Logical steps in active inference, following Peirce [103] *Source* [16]

Inference and Hypothesis

- Deductive or Analytic
- Synthetic
  - Induction
  - Hypothesis/Abduction in context & ecology

Fallibilism, a term frequently used by Peirce, was paramount to this approach. This is not surprising, because he also understood something about the new tools that were emerging in the statistical sciences, such as the properties of chance and probability that involved looking for frequencies, determining statistical support, and flushing out the controls and the control groups from experimental design [5, 7, 29].

Importantly, for [7], pragmatism is knotted to "abduction"; abductive events are at the heart of inquiry, abduction is about hypothesis creation. Abduction, in other worlds, is about hypothesis formation and testing; the grounding is in objects and action, reference and a natural sense of objects. The richer the hypothesis, in some contexts the larger the implication: the deeper the hypothesis, the better the understanding, inferences towards better explanations, but abduction is mostly about good enough fit for the context of epistemic satisfaction [30].

A logic of discovery is not one set of epistemic events (c.f. [22, 31, 32]). A logic of discovery as [33] would suggest is a "manual of conclusions", but an assessment of inferences made, the problematic in a particular context, the fit of the hypothesis and the explanatory depth or not; but set in a context of effective strategies, research trends of success, and reliable justification of inference rules (see also [34]), constrained and situated in ecological contexts [35, 36]. Abduction in context; options, suitability and viability; testing and foraging are continuous with these events.

What is an instinctive adaptation; they are ideas that fit easily into a niche, where there is probably selection pressure, that allows for a good enough fit. One orientation of the living brain is statistical inference, the flip side of which is probabilistic errors and to minimize error [37] and energy expenditure by the brain [38].

The theory of inquiry, of hypothesis testing, is rooted in this cognitive perspective; fixation of belief is rooted in the organization of action [13, 39]. The orientation is not simply reactive, but anticipatory—but also responsive to discrepancy with expectations. Indeed [40], a scholar of Peirce has noted that Kant and Alexander Bain were particular influential with regard to the origins of pragmatism. Alexander Bain, less well known, had written an influential book in the early nineteenth century on the will, linking beliefs to a tendency and strength of action. And abduction is an inference that emerges within this context, generating a new idea, slightly altered one, etc. All of which is tied to foraging for epistemic coherence.

Of course, action is often habitual routine. Memory, attention, and other cognitive resources are minimized during diverse routines; cognitive capabilities are then recruited elsewhere in the ongoing action. It is the breakdown that helps generate further action and cognitive resources to learn, attend, and construe new resolutions and new forms of adaptations, and part of the clarity to determine the diverse consequences of the new set of hypotheses determinant in human action, long term [7, 16, 29].

2 Keeping Track, Abduction and the Predictive Brain

Ecological Considerations: Cognitive adaptations towards objects that makes their ready use and understanding available, a veritable set of cephalic heuristics [5, 41, 42] we arrive at some ideas rather quickly. We arrive at them because there is a background context of adaptation; the ecological fluidity to the living brain oxygenated the ease of hypothesis formation or abductive inference [3].

Anchoring our decision making to useful heuristics with a cephalic orientation about number and frequencies is very useful. The way problems are framed influences the judgment, the decision making. Keeping track is demythologizing the decision making by making it transparent and readily understandable; no easy chore, but knowing something about relationships between events, the origins of statistics being linked to keeping track of events [43]. One way we do this is by representing statistical relationships and the quantification of data in terms of probabilities and natural frequencies. Keeping track of events is expanded by the scaffolding to diverse forms of information [44, 45].

Abduction in context, constrained by context and ecology. Our sensory capacity is keen to detect objects that afford sustenance or harm [5, 46]. Fast forms for detecting information can change the structure; diverse forms of heuristics, fast ways to solve problems both specific and general, evolved in the evolution of our brain.

All inquiry constrained by frameworks utilizing diverse forms of statistical inferences and underlying the logics of discovery are very Peircian indeed, and quite modern. And his epistemology was emboldened less by the issue of correspondence than coherence and the fruits of labor in inquiry.

We come prepared to associate a number of events linked by causal building blocks in cephalic structures by worldly events. *Ecological rationality*—viz., readily available heuristics well-grounded in successful decision making [42]—places decision making and the use of statistical features within cephalic predilections about numbers and representations of frequencies in real contexts. Gerd Gigerenzer [5] following Herbert Simon looked for ways in which we can adapt and satisfy ourselves with the best we can discern given the context, circumstances and expectations in behavioral explanations. There are better and worse forms of decision making.

As Peirce noted, we come with a predilection to hit on the right hypothesis, and to react quickly to solve problems, both specific and general: diverse heuristics developed in the evolution of our brain. In other words, we come prepared to associate a number of events linked by causal building blocks [13, 15, 47]. Ecological sensibility and rationality, abduction in context describe readily available heuristics well-grounded in successful decision making [23, 48]. Abduction is grounded in ecological fluidity and social evolution [49], availability of resources and neural capability [50].

Inquiry and learning is in a context when expectations are not met, when habit formation and activity are disrupted [6, 7, 51]. Prediction is coupled with expectancies and their breakdown [52, 53]. When expectations are thwarted, a broad array of learning occurs through new problem solving and search principles. Information acquisition and control is a strong desire. Aversion to ambiguity, for example, is a real property of our decision-making. Ambiguity can breed indecision, the status quo in action in human decision-making processes [54]. Foraging for coherence and sampling environmental stimuli of significance is tied to the fact that, within traditional learning, associations are embedded within probability of occurrences. Learning, abduction, in part, takes place when expectations and coherent forms of behavioral adaptation are compromised or called into question [55].

When expectations are thwarted, a broad array of learning occurs through new problem solving and search principles. Peirce's view of inquiry and learning was prescient, for the variants of this view would capture learning theory through what became known as the [55] equation:

$$\Delta V = \alpha\beta(\lambda - V)$$

The Rescorla-Wagner model depicts the associative strengths of stimuli and how discrepancies from expectations are resolved. An association, and thereby learning, occurs by the strength of the predictions that are being developed. The model then is not simply a mathematical approach to neural science, but also an incorporation of a cognitive point of view. In the equation, V represents the current associative strength of the stimulus, while λ shows the maximum associative strength of the primary motivating event.

Abductive inference and the detection of salience in the context of hypothesis formation and habit generation when expectations are thwarted, when the anticipatory behaviors have lost credence, and when the expected outcomes are undermined. From there comes a further foraging for coherence and propagating actions that are consistent with newly formed hypotheses and expectations, abduction is set in this context; the search that underlies appetitive search to remedy a problematic context that generates epistemic motion [3, 51].

Much of real inquiry, however, involves struggling with the not-known, the barely-known, or the yet-to-be-known [51]. Indeed, from our routine problem-solving for survival emerged our ability to discern and act on that which is knowable only with degrees of confidence. That judgmental practice is the heart of statistical inference.

We now know that diverse regions of the brain are linked to statistical inference, and they include both neocortical and subcortical regions such as the basal ganglia [56]. Perhaps it is not surprising that cortical regions are linked to statistical inferences, but the subcortical sites may be surprising, as these sites are also understood in the context of motor regions. The basal ganglia is the head ganglia of motor control, but then movement and action are inherently related in the organization of action and in the anticipatory regulation of sensory and motor expression and expectations [57, 58].

Resolving uncertainty clearly is a major motivator of behavior. Perhaps this is what makes us more eager to hear the end of a song that defies our expectations rather than fulfilling them, but the brain is wired to minimize neural exhaustion [38], to minimize neural excess use.

Foraging for coherence is active sampling tied to prediction of action and expectations of events. Typically, it runs continuously. Of course, our world is mostly coping, managing and surviving, with bouts of relief and neural regeneration [38, 59]. Breakdowns result in abductive inference or active inference and learning [50].

Interestingly, information deprivation is construed as a cognitive deprivation. A consequent hunger to fill this gap occurs [60]. Within the bounds of reason, some forms of inquiry and cognitive coherence may reflect this search for cognitive equilibrium. The emphasis is on embodied knowing; visceral input is something vital to curiosity, learning, and inquiry. Interestingly, in the common phrase, “to pique one’s curiosity,” the word pique means “to arouse or provoke,” but the word can also mean “to cause to feel vexation or resentment.” The gap of information, both relative and absolute, leads to a possible understanding of curiosity and the joys that one might find in the search for and attainment of information. The relationship of the information gap to curiosity has been articulated through the use of information theory’s entropy coefficient:

$$\sum_{\mathrm{p_i}}^{n} \log_2 \mathrm{p_i i} = 1$$

In this equation, *n* represents the total possible choices or outcomes of information, whereas pi is the assessed probability that a particular choice will occur. As knowledge concerning each choice increases, the probabilities of each become more varied and exact. The equation is useful, but not necessarily exact, in quantifying multidimensional information in a one-dimensional manner. In terms of the information gap and curiosity levels, several entropy measures are necessary: the individual’s current situation, the informational goal of the individual, and a situational level of ignorance are all factors. The absolute magnitude would therefore be the informational goal minus the current situation. The relative magnitude would be the absolute magnitude divided by the difference of the informational goal and level of total ignorance. Abduction emerges in filling in the gaps, the active inferences that underlie the curiosity, the discomfort, the noted discrepancy, and ameliorates neural energy consumption.

Abduction and Surprise: Neuroscientific inquiry, in the same way, is geared to minimizing surprises through investigation. The neural orientation is to gain precision in perceptual contact with minimizing neural energy expenditure [38]. Conservation of energy one principle adaptation [61]. Importantly, the view is tied to classical pragmatism, in the form of [6, 7, 16, 62] and particularly [17, 51].

Effort and the engagement of the motor system is inherently tagging and predicting events with contact with real or imagined object, whether conscious or not, and perhaps mostly not [15]. This may, perhaps, occur in particular during default modes of simulation [63] or simply be a running neural capability in suitable contexts. Prediction of events is the running theme in a living brain tied to expectations and regulation [64]. Diverse regions of the cortex [65] and subcortical regions

[58] underlie the action sequences, and the abductive inferences that underlie the organization of action.

The “pragmatic turn in cognitive science” [66] recognizes the continuous organization of cognitive/motor regulation of action. Diverse forms of dualism are diluted [17], such as thought and action, perception and thought, in the organization of action. As [17] noted “there is simply a continuously ordered sequence of events”. And that is essential in the continuous integration of afferent and efferent timing of events and in the predictive capability that runs through the motor systems and the generation of action sequences.

One dominant theme in foraging for coherence is the detection of error and the use of various methods to correct effort prediction. Expectations and feedback are a dominant and recurrent theme across cephalic systems and capabilities. These computational systems, embedded in well—worn and tested practices, cultural contexts at the heart of finding meaning across diverse systems, particularly motor/action/appraisals—a key feature of cephalic adaptive systems [5] and adaptive abductive inferences.

2.1 Bayes’ Theorem and Bayesian Brain: Abductive Inference

This orientation to self-correct and keep track of prior probabilities is a key feature of the brain [54]. Bayes’ theorem is one way in which to understand expectations in terms of prior probabilities that are embedded in our predictions about recurrent events, and learning from them. These events lie within a wide array of problem-solving capabilities, without one set of consistent overlying rules (see also [15, 30]). In other words, no one unifying super problem-solving device exists; rather, we have a wide assortment of adaptive tools that underlie problems involving the evolution of adaptive systems and abductive inference.

Abductive inferences is within self-corrective neural systems. Bayes’ theorem is a fundamental theorem of probability that states that, for any two events A and B, the probability of A given B can be computed from the probability of B given A, as well as the overall probabilities (known as the “prior probabilities”) of A and B.

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

Moreover, we may be prepared to understand natural frequencies more easily than other ways of representing events as we track objects and events with diverse heuristic devices and revise our orientations to events if we need to.

A Bayesian cephalic self-corrective, a realistic metaphor, functions to coordinate expectations with new forms of evidence in a context of predictive capabilities [67]. This is perhaps one part of cephalic capabilities—a ‘grab bag’ of diverse functions

[15] that underlie human foraging for coherence and sampling behaviors and adaptive good enough and contextual adaptive inferences.

2.2 Cataloging; Conditions for Abductive Inference

But one thing that may be inherent in cephalic function and adaptative systems are probabilistic expectations of diverse forms of sensory information and ways in which to track, link and structure events in meaningful categorical relationships that underlie causal inferences [68].

It is not surprising that we come prepared to catalogue and group objects, especially into things alive or not-alive. The epistemic attempts of Linnaeus (and others, such as Buffon) represent just one human scheme for classifying plants and animals; human beings have always explored their world through categories of understanding.

Indeed, categories of understanding—classificatory systems embedded in projectable predicates [26] and with expectations—are part of our essential tagging capabilities. We do this easily; it is fundamental to inductive events and for tracking and tagging objects of relevance in categories that match objects to predictable statistical relationships, perhaps using Bayesian type tracking and sampling mechanisms, namely building on prior probabilities through evidence and continuous updating as needed [68].

This is a valuable tool in a world of adaptation and heuristic problem-solving [5, 30]. The natural continuum of events is from theoretical constructs to solutions that attempt to satisfy very real problems. The "naturalization of intelligence," a phrase Dewey used, is to be always oriented to the problem. It begins in organizing events into categories of coherence with an eye towards things and objects that matter to us or which we wish to avoid [27].

Natural knowledge is an ancient predilection for taming reality by categories, a skill we come prepared with and which is linked to diverse forms of cephalic capability. Coming prepared to recognize, for instance, edible objects from inedible ones, anchors the way forward in a suitable environment in which additional learning can take place. And, of course, having capability to draw adaptive abductive inferences.

Human beings have a common currency of survival and folk biological categories. These are shared categories of understanding in common adaptation while we forage for both survival and understanding. Indeed, the two mix in an endless stream as we sample and anticipate events to come while we sustain our viability. Inferential systems are grounded in our sense of nature: the evolutionary equipment that got us here some 10,000 years ago [69].

Indeed, exposure to diverse forms of visual experience figures in both visual recognition and the semantic meaning of animate objects, e.g., plant recognition [4].

Developing postnatal brains are formed by exposure to events. The sculpting of diverse brain regions is a continuous function during development, leading to both pre- and postnatal achievement of basic form and function.

The search is for the "stable amidst the precarious" [51] in a never-ending battle for viability. Classification is rooted in that core search. Theoretical taxonomies are always operative. Without theory, there is no lens, no seeing, for coherence [22]. Inferences are embedded in an orientation of action and in viability of outcomes, expectations in past habits codified by neural systems within environmental constraints on hypothesis formation.

Our commitment to the taxonomy of objects is built on the success of that taxonomy as it renders events predictable and coherent. It is not make-believe. The whole point about both pragmatism and adaptation is that things matter, there are consequences, and the consequences are not abstract. That is the allure of pragmatism—a pragmatism grounded in discovery and understanding. Such taxonomic cultural evolution underlies the evolution of knowledge. We look for core features of objects that provide us with some coherence in expectations.

Abductive inference is set in a context or prior order, whether as part of neural core function or as selected objects and groupings. Core concepts include that of agency, causation, time space, number, and our sense of objects. Discerning recurrence and reliability in the network of semantic meaning is a "good enough" way of making sense of our world.

However, a set of cognitive capabilities to recognize the usefulness of natural objects varies with culture. Atran and Medin, for instance, in comparing the ability of the Itza people versus local non-natives in picking out natural objects, reveal something like devolution of function [70]. Folk biological knowledge is in a context, even in its most primitive modern form. What this sort of finding might be saying is that we are prepared to note useful objects. They figure in reliable information in the organization of action and decision making within a culture. Indeed, categories are endlessly fallible. Peirce's fallibilism, on the other hand, sets expectations on limits.

The platypus aptly illustrates some of the limits of our sampling systems. The main definition of a mammal is an animal that bears live young (i.e., does not lay eggs). Of course, a platypus is just that—an egg laying mammal—a marvel of nature and a delight to behold, but also something both within and outside our classificatory system. The point here is about seeing, context, classification, and the expansion of knowledge, both operative when foraging for coherence and updating by sampling continuously. Abductive connects the dots, despite the discrepancy.

Aids for Abductive Inference: A wide array of both specific and more general problem-solving capabilities is vital in foraging for coherence; certainly, numerical capability is essential for tracking events. Spatial capability figures in noting where things are and when they might appear, and what events are associated by predictability or non-predictability.

A tool box in epistemic discoveries, abductive inferences is tied to core knowledge about space, time, probability, agency, language, statistical capability, and a variety of social and technological skills. We imagine the preconditions for foraging for coherence and then place them in a sense of inquiry with constraints on hypothesis formation and inferences towards good explanation [7, 16], or a good enough explanation, given the context (see also [30]).

Reason and problem-solving aim toward a “good fit,” constrained by ecological context and evolutionary history (e.g. [5, 30]). Then we have something Peirce might have expressed, if we go further in integrating a biological perspective with hypothesis formation and abduction. We come prepared with diverse heuristics in problem-solving. Heuristics—i.e., fast ways to solve problems both specific and general—evolved along with our brain.

Abductive inference is in context, constrained by context and ecology. Our sensory capacity is keen to detect objects that afford sustenance or harm [71]. Fast forms for detecting information can change the structure. Indeed, we come prepared to associate a number of events linked by causal building blocks in cephalic structures by worldly events [15, 47]. *Ecological rationality*—readily available heuristics well-grounded in successful decision making, as the classical pragmatists understood—places decision making and the use of statistical features within cephalic predilections about numbers and representations of frequencies in real contexts [5].

3 Conclusion: Abduction and Inquiry

All inquiries constrained by frameworks utilizing diverse forms of statistical inferences and underlying the logics of discovery are very Peircean indeed and extend into modern times [25, 43, 72]). The understanding mind requires thinking about our evolution, our problem-solving capabilities, our ability to predict events, and our sense of probable events in space and time as tied to habits of intelligence and problem solving. Indeed, keeping track of objects is a basic predilection built into cephalic sensibility and reaching into everything we do.

The intellectual move is away from excessive essentialism and rationalism, while being anchored to objects within a sense of cognitive adaptation. The philosophic sensibility is oriented to action. And the brain is oriented to maximize good enough abductive inference to resolve problematic contexts and the breakdown of expectations. Epistemic and existential coherence are anchored to heuristics that underlie abductive inference; hitting on the right hypothesis in contexts of meaning and local adaptation.

Epigenetic regulation of diverse regions of the brain underlie inquiry or learning vital for life [73]. In other words, regions of the brain are modified by experience. While other animals have some plasticity or neurogenesis, in humans it is vast: look at the diverse languages we can learn, the many forms of epistemic and other human activities we can express, the different inventions that we can generate, that expand our sense for seeing, hearing, and knowing, that underlie abduction.

Foraging for coherence is grounded in action, in never losing sight that we are object/body oriented. Despite the perplexities of inquiry, this is a primary anchoring point (Quine 1951). We traverse an epistemic space of endlessly boot-strapping to get a grasp on predictive capability. The issue is coherence and reliability for projectable predicates of meaning [26]. The epistemic foraging metaphor that underlies abductive inferences is within a framework in which the more reliable is more heavily embedded in our frames of understanding. A movement from hands to language is a continuous function of our communicative sensibility; gestural expression predominates as we explore, cooperate, and communicate with others [74].

Context is everything. The brain includes both the central and peripheral systems and is deeply embedded in social milieu. Take memory, for example: memory is not exactly in the brain, but extends into the environment. Indeed, a number of investigators have detailed ways in which memory and other cognitive abilities are endlessly tied to the environment in which we are foraging for coherence [75].

Action and perception are linked as we explore the world [17, 37]; our sense of ourselves, imagined or not, is knotted to action [75]. There are many ways to talk about the mind not being strictly in the head, but rather as extended across the terrain. In this terrain, we are foraging for coherence with cephalic systems prepared to understand objects easily. For example, we recognize faces and places, distinguish animate from inanimate objects, observe movement, and infer causal relationships early in development and a feature across a lifetime of problem-solving, etc. Social contact is essential for us, and is tied to the actions of others, their direction, and their movement.

Abduction is set in the context of others; what others have acquired. The social roots permeate the abductive space of decision making. The brain is not isolated, and neither are we as a species. We are embedded in the culture, in the local niche, in the diverse practices in which we participate. Understanding cephalic systems leads us to acknowledge this profound fact. We are inherently in a universe of coherence; chaos and utter breakdown are actually incredibly rare.

Coherence, however, is not the same as truth-seeking and self-corrective inquiry; we must never confuse these concepts. Sampling and foraging for coherence are in a context of inquiry, of checking and modifying habitual forms of behavior. This is not de novo. The practices in which we participate are well-orchestrated and pervade the organization of action [75]. We come into the world prepared to recognize objects as meaningful or not, animate or not, reliable or not, etc.; then we are embedded in diverse practices that pervade our coherent forms of action [76–102].

Human neural organization and the stability of habit formation and continued viability is based largely on anticipatory cognitive system tied to ecological contexts of relevance and neural design. For Peirce, pragmatism is knotted inextricably to *abduction*. By abduction, Peirce meant the genesis of ideas or, in other words, hypothesis formation and testing. Abductive events are at the heart of inquiry. For Peirce, abduction is “preparatory;” and induction is a “concluding step” grounded in natural outcomes and deduction. By introducing the idea of abduction, Peirce offered an alternative to rationalism and deduction, as well as endless empiricism and induction.

In conclusion, predictive predilections knotted to abduction or hypothesis testing dominates the landscape of diverse forms of problem solving. Abduction is biologically constrained and contextual, not a monolithic term and runs the range of neural capability. We come oriented to exclude some forms of irrelevant forms of information in problem solving that aids problem solving and predictive coherence. Abduction from the mundane and the adaptive to the exploratory is reflective of active problem solving. Indeed, abduction rest on the fact that brain is an active organ in which plasticity of function or neurogenesis is as basic as neural or synaptic pruning.

References

1. Schulkin, J.: Social Sensibility and Neural Function. MIT Press, Cambridge (2000)
2. Schulkin, J.: Bodily Sensibility Intelligent Action. Oxford University Press, Oxford (2004)
3. Schulkin, J.: Foraging for Coherence in Neuroscience. Palgrave Macmillan, London (2015)
4. Carey, S.: On the origins of concepts. Oxford University Press, Oxford, United Kingdom (2009)
5. Gigerenzer, G.: Adaptive thinking. Oxford University, Oxford, United Kingdom (2000)
6. Peirce, C.S.: The fixation of belief. Popular Science Monthly **12**, 1–15 (1877)
7. Peirce, C.S.: Deduction, induction and hypothesis. Popular Sci. Mon. **13**, 470–482 (1878)
8. Lakoff, G., Johnson, M.: Philosophy in the Flesh. Basic Books, New York (1999)
9. Schulkin, J.: Cognitive Adaptation: A Pragmatist Perspective. Cambridge University Press, Cambridge (2009)
10. Dennett, D.: The Intentional Stance. MIT Press, Cambridge, Massachusetts (1987)
11. Kitcher, P.: Preludes to Pragmatism. Oxford University Press, Oxford (2012)
12. Dewey, J.: The Influence of Darwin on Philosophy. Indiana University Press, Bloomington (1910, 1965)
13. Gallistel, C.R.: The Organization of Learning. MIT Press, Cambridge, Massachusetts (1990)
14. Keil, F.: Concepts, Kinds and Cognitive Development. MIT Press, Cambridge (1979,1992)
15. Rozin, P.: The evolution of intelligence and access to the cognitive unconscious. In: Sprague, J., Epstein, A.N. (eds.) Progress in Psychobiology and Physiological Psychology, vol. 6, pp. 245–280. Academic Press, New York, NY (1976)
16. Peirce, C.S.: Questions concerning certain faculties claimed for man. In: Houser, N., Kloesel, C. (eds.) The Essential Peirce, vol. I, pp. 11–27. Indiana University Press, Bloomington, Indiana (1868, 1992)
17. Dewey, J.: The reflex arc concept in psychology. Psychol. Rev. **3**, 357–370 (1896)
18. Parrott, G.W., Schulkin, J.: Neuropsychology and the cognitive nature of emotions. Cogn. Emot. **7**, 43–59 (1993)
19. Houser, N., Roberts, D.D., Evra, Van: Studies in the logic of Charles Sanders Peirce. University of Indiana Press, Indianapolis, Indiana (1997)
20. Houser, N., Eller, J.R., Lewis, A.C., De Tienne, A., Clark, C.L., Bront Davis, D.: The essential Peirce, vol. 2. Indiana University Press, Bloomington, Indiana (1998)
21. Anderson, D.R.: Science of Discovery. Purdue University Press, West Lafayette, Indiana (1995)
22. Hanson, N.R.: Patterns of Discovery. Cambridge University Press, Cambridge, United Kingdom (1958, 1972)
23. Heelan, P.A., Schulkin, J.: Hermeneutical philosophy and pragmatism: a philosophy of science. Synthese **115**, 269–302 (1998)
24. Wittgenstein, L.: Philosophical Investigations. Macmillan Publishing, New York, NY (1953, 1968)
25. Hanson, N.R.: Observation and explanation. Harper Press, New York, NY (1971)
26. Goodman, N.: Fact, Fiction and Forecast. Bobbs-Merrill, New York (1955, 1973)
27. Altran, S.: Cognitive Foundations of Natural Philosophy. Cambridge University Press, Cambridge (1990)
28. Hohwy, J.: The Predictive Mind. Oxford University Press, Oxford (2013)
29. Peirce, C.S.: The architecture of theories. Monist **1**, 61–76 (1892)
30. Simon, H.A.: Models of Bounded Rationality. MIT Press, Cambridge (1982)
31. Hintikka, J.: What is abduction? Trans. CS Peirce Soc. **34**, 503–533 (1998)

32. Rescher, N.: Essays in the History of Philosophy. Aldersot, Averbury (1995)
33. Hanson, N.R.: Notes toward a logic of discovery. In: Bernstein, R.J. (ed.) Perspectives on Peirce. Yale University Press, New Haven (1965)
34. Paavola, S.: Abduction cognition: the epistemological and eco-cognitive dimensions of hypothetical reasoning. Trans. CS Peirce **47**, 252–256 (2011)
35. Magnani, L.: The abductive structure of scientific creativity: an essay on the ecology of cognition. Springer, Berlin (2017)
36. Shook, J.R.: Abduction, complex inferences, and emergent heuristics of scientific inquiry. Axiomathes **26**, 157–186 (2016)
37. Friston, K.: The free-energy principle. Cell **30**, 781–899 (2009)
38. Peters, A., McEwen, B.S., Friston, K.: Uncertainty and stress. Prog. Neurobiolgy **156**, 164–188l (2017)
39. Gallistel, C.R.: The organization of action: A New Synthesis. Lawrence Erlbaum, Hillsdale, New Jersey (1980)
40. Fisch, M.H.: Peirce, Semiotic and Pragmatism. Indiana University Press, Bloomington, IN (1986)
41. Brighton, H., Gigerenzer, G.: Homo heuristicus and the bias-variance dilemma. In: Schulkin, J. (ed.) Action, Perception and the Brain. Palgrave Macmillan, London (2012)
42. Goldstein, D.G., Gigerenzer, G.: Models of ecological rationality: the recognition heuristic. Psychol. Rev. **109**, 75–90 (2002)
43. Hacking, I.: The Taming of Chance. Cambridge University Press, Cambridge, United Kingdom (1999)
44. Clark, A.: Being There. MIT Press, Cambridge, Massachusetts (1997)
45. Donald, M.: A Mind So Rare: The Evolution of Human Consciousness. Norton, New York, NY (2001)
46. Kahneman, D., Slovic, P., Tversky, A.: Judgment under uncertainty: heuristics and biases. Cambridge University Press, New York, NY (1982)
47. Rozin, P.: Evolution and development of brains and cultures. In: Gazzaniga, M.S., Altman, J.S. (eds.) Brain and mind: Evolutionary perspectives, pp. 111–123. Human Frontiers Sciences Program, Strasbourg, France (1998)
48. Schulkin, J.: Naturalism and Pragmatism. Palgrave Macmillan, London (2012)
49. Kirk, A., Stokerson, P.: Naturalizing Peirce's semiotics: ecological psychology's solution to the problem of creative abduction (2010)
50. Friston, K., et al.: Active Inference and Learning. Neurosci. Biobehav. Rev. **68**, 862–869 (2016)
51. Dewey, J.: Experience and Nature. Open Court, La Salle, Illinois (1925, 1989)
52. Miller, N.E.: Liberalization of basic S-R concepts: extensions to conflict behavior, motivation and social learning. In: Koch, S. (ed.) Psychology: A Study of a Science, vol. 2, pp. 198–292. McGraw-Hill, New York, NY (1959)
53. Rescorla, R.A.: Pavlovian conditioning. Am. Psychol. **43**, 151–160 (1988)
54. Baron, J.: Thinking and Deciding. Cambridge University Press, Cambridge, United Kingdom (1988, 2008)
55. Rescorla, R.A, Wagner, A. A theory of pavlovian conditioning. In: Classical Conditioning. NY: Apleton-Century Crofts (1972)
56. Glimcher, P.W.: Decisions, Uncertainty and the Brain. MIT Press, Cambridge (2003)
57. Jeannerod, M.: Motor Cognition. Oxford University Press, Oxford (2006)
58. Schulkin, J.: Effort: A Behavioral Neuroscience Perspective on the Will. Erlbaum Press, Mahjaw, NJ (2007)
59. Schulkin, J.: Uncovering An Information Molecule: Corticotrophin Releasing Factor. Oxford University Press, Oxford (2017)
60. Loewenstein, G.: The psychology of curiosity. Psychol. Bull. **116**, 75–98 (1994)
61. Helmholtz, H.: On the conservation of force. In: Science and Culture: Popular and Philosophical Essay: Herman von Helmholtz. University of Chicago Press, Chicago (1862–1863, 1995)

62. James, W.: The Principles of Psychology, vols. 1–2. Henry Holt, New York, NY (1890, 1952)
63. Schubotz, R.I.: Prediction of external events with our motor systems. Trends Cogn. Sci. **11**, 211–218 (2007)
64. Bubic, A., Cramon, D.Y., Schubotz, R.I.: Prediction, cognition and the brain. Front. Hum. Neurosc. **4**, 22–44 (2010)
65. Cisek, P.: Cortical mechanisms of action selection: the affordance competition hypothesis. Philosop. Trans. Royal Soc. **B12**, 2032–2044 (2007)
66. Engel, A.K., Kriston, K.J., Kragic, D.: The Pragmatic Turn. MIT Press, Cambridge (2015)
67. Glymour, C: Thinking Through Things. MIT Press, Cambridge (1992, 2004)
68. Tennenbaum, J.B., et al.: How to grow a mind: statistics, structure and abstraction. Science **331**, 1279–1285 (2011)
69. Mellars, P.: Why did modern human populations disperse from Africia ca. 60,000 years ago. Proc. Nat. Acad. Sci. **103**, 9381–9386 (2006)
70. Altran, S., Medin, D.: The Native Mind and the Social Construction of Nature. MIT Press, Cambridge (2008)
71. Gibson, J.J.: The Senses Considered as Perceptual Systems. Houghton-Mifflin, New York (1966)
72. Hacking, I.: Logic of statistical inference. Cambridge University Press, Cambridge, United Kingdom (1964)
73. Waddell, J., Shors, T.J.: Neurogenesis, learning and associative strength. Eur. J. Neurosci. **27**, 3020–3028 (2008)
74. Corballis, M.C.: From Hand to Mouth. Princeton University Press, Princeton (2002)
75. Noë, A.: Action in perception. MIT Press, Cambridge, Massachusetts (2004)
76. Clark, A.: Surfing Uncertainty. Oxford University Press, Oxford (2016)
77. Descartes, R.: Discourse on Method. Hackett Publishing Co., Indianapolis, Indiana (1637, 1993)
78. Dewey, J.: Logic: The Theory of Inquiry. Holt, Rinehart and Winston, New York (1938)
79. Hacking, I.: The emergence of probability. Cambridge University Press, Cambridge, United Kingdom (1975)
80. Hoffman, M.: Problems with Peirce's concept of abduction. Found. Sci. **3**, 271–305 (1999)
81. Houser, N., Kloesel, C.: The Essential Peirce (Vol. I). Indiana University Press, Bloomington, Indiana (1992, 1998)
82. Kahneman, D.: Thinking Fast and Slow. Farrar, Strauss and Giroux, New York (2011)
83. Keynes, J.M.: A Treatise on Probability. Harper and Row, New York, NY (1921, 1957)
84. McEwen, B.S., et al.: Sixty years of neuroendocrinology. J. Endocrinol. **6**, 226–267 (2015)
85. Nickles, T.: Scientific Discovery, Logic and Rationality. Reidel, Dordrecht (1980)
86. Paavola, S.: Abduction as a logic and methodology of discovery. Found. Sci. **9**, 267–273 (2004)
87. Peirce, C.S.: Studies in logic by members of the Johns Hopkins University. Little, Brown and Co, Boston, Massachusetts (1883)
88. Peirce, C.S.: On a method for search for categories. In: Houser, N., Kloesel, C. (eds.) The Essential Peirce, vol. I. Indiana University Press, Bloomington, Indiana (1866)
103. Peirce, C.S.: Deduction, induction and hypothesis. Popular Scientific Monthly **13**, 470–482 (1878)
89. Peirce, C.S.: Reasoning and the Logic of Things. Harvard University Press, Cambridge, Massachusetts (1899, 1992)
90. Peirce, C.S., Jastrow, J.: On small differences of sensation. Memoirs Nat. Acad. Sci. **3**, 75–83 (1885)
91. Pezzulo G.: Anticipation and the predictive brain
92. Quine, W.V.O.: From a Logical Point of View. Harper and Row, New York (1953, 1961)
93. Quine, W.V.O.: Ontological Relativity and Other Essays. Columbia University Press, NY (1969)
94. Reid, T. Essays on the Intellectual Powers of Man. MIT Press, Cambridge (1785,1969)

95. Shanahan, T.: The first moments of scientific inquiry: CS Peirce on the logic of abduction. Trans. CS Peirce Soc. **22** (1986)
96. Swanson. L.W.: Brain Architecture. Oxford University Press, Oxford (2003, 2013)
97. Thagard, P.: Coherence in Thought and Action. MIT Press, Cambridge (2000)
98. Thagard, P. Abductive inference: from philosophical analysis to neural mechanisms. In: Inductive Reasoning. Cambridge University Press, Cambridge (2007)
99. Turing, A.: The Essential Turing. Clarendon Press, Oxford, United Kingdom (2004)
100. Van Frassen, B.S.: The Empirical Stance. Yale University Press, New Haven (2002)
101. Van Essen, D.C., et al.: Information processing in the primate visual system. Science **25**, 419–422 (1992)
102. Weiss, P.: Charles S. Peirce: Philosopher. In: Bernstein, R.J. (ed.) Perspectives on Peirce: Critical essays on Charles Sanders Peirce, pp. 139–148. New Haven, Connecticut: Yale University Press (1965)

Abductive Inference, Self-Knowledge, and the Myth of Introspection

Eric Phillip Charles and Nicholas S. Thompson

Abstract Much of the history of psychology can be understood as a debate over what we do when we attribute psychological states to ourselves and to others. In the classic Cartesian view, those activities are quite distinct: We engage an infallible ability—introspection—when examining our own psychological states, but merely speculate when trying to identify the psychological states of others. The American Philosopher, Charles Sanders Peirce dedicated several early papers to a critique of the Cartesian approach. He concluded that attempts at self-knowledge require the same inferential processes that we use when attributing mental states to others, and therefore incur the same logical risks. By pursing these ideas further, we intend to show that self-knowledge results from a special kind of abduction; the inference of behavior states from particular observed behaviors. Such inferences allow us to anticipate yet-unseen patterns of behavior in yet-to-be-manifested circumstances.

1 An Alternative to Cartesian Introspection

Questions about the nature of self-knowledge date back to the earliest days of philosophy. If we look to the ancient world, we find the instructions to "Know Thyself" (Σαυτὸν ἴσθι) written over the entrance to the Temple of Apollo at Delphi, where the great Oracle resided. We may presume, because many classical philosophers elaborated upon those instructions, that Greeks considered self-knowledge a difficult, but attainable, goal. The idea of self-knowledge as a significant achievement contrasts sharply with Cartesian dualism, which convinced generations of philosophers that self-knowledge was the foundation of all knowledge. By the latter way of thinking, knowledge about self is not only in-practice easier to obtain

E. P. Charles (✉)
American University, Washington, USA
e-mail: echarles@american.edu; Charles.Psychology@gmail.com

N. S. Thompson
Clark University, Worcester, USA
e-mail: Thompnickson2@gmail.com; nthompson@clarku.edu

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_12

than knowledge about others, it is easier to obtain in principle, as the two types of knowledge are fundamentally different. Although this Cartesian conviction has become more and more controversial, much psychological research and theory still proceed on the assumption that self-knowledge is relatively easily obtained, whereas knowledge-about-the-minds-of-others is harder to obtain.

In psychology, the continuing influence of Descartes is readily seen in the vast "theory of mind" literature, in which empirical researchers theorize about what we do when "attributing" mental states to others [9, 14, 21]. In one way or another, those researchers assert an introspective step, whereby we look into ourselves, and then a projective step, whereby we mentally-project that self-knowledge onto another being. Proponents of that approach assume—without question—that the introspective step entails a relatively direct experience of one's own mind, contrasted with guesses about the minds of other people; they then attempt to research how we get better at such guesses over time. These approaches, building out of the Cartesian tradition, take it as foundational that the minds of others are inherently not the types of things that we can experience [2, 3, 13].

The status of dualistic thinking in psychology is often hidden behind authors' adamantly denying being "dualists," while still advancing inherently dualistic ideas. Most of our research-psychologist colleagues would be embarrassed to assert that mind and matter are different substances, yet some are happy to refer to "inner" causes located in some "deep" place within "the mind." As a professional matter, the research psychologists we are acquainted with recognize that mind-world and mind-body dualisms are broadly problematic. If pressed they will fall back on an amorphous brain-state materialism, claiming that anything about their theory that seems dualistic is shorthand for some sort of unspecified neuronal process. The commonness of this weak defense—embracing the language and concepts of mind-world dualism while pleading for an incompatible materialism—is due, in part, to the lack of clear alternatives to the dualistic language and concepts [6, 7]. We have shown elsewhere that such claims entail circular reasoning [21].

One alternative to such dualism is the pragmatism of Charles Sanders Peirce. We have elsewhere argued that Peirce's work (implicitly) underlies several threads in twentieth century psychology, including branches of behaviorism and ecological psychology [4, 5, 7, 24], and others have made similar connections [e.g., 8, 11, 20]. In that context, we have been working to develop principles regarding psychology that are grounded in Peirce's early work. These ideas will be further developed here, via an exploration of the problem of introspection and abductive inference.

2 Early Peirce: Cognition as Inference

Four of Peirce's early ideas will be the most crucial in the analysis that follows: (1) The rejection of a capacity of "intuitive" introspection; (2) the process of inference by abduction; (3) clarifying meaning using the pragmatic maxim; and (4) Peirce's definition of "belief." Once those ideas are in hand, we will use them to clarify our thinking about self-knowledge.

2.1 Incapacities of Man

Several of Peirce's early philosophical works focused on rejecting the existence of certain mental "capacities" that past philosophers (particularly in the Cartesian tradition), believed were possessed by all people (e.g., [15–17]) In those works Peirce repeatedly argued, contra Descartes, that people do not have a special intuition of introspection. In this context, "intuition" indicated a cognition that did not result from inference ([15], including footnote 1). Peirce pointed out that there was no evidence for such an intuition, and argued instead that self-knowledge is attained via an inferential process of the same kind as our inferences about the minds of others. Thus, the supposed results of introspective intuition could not be understood "except as a hypothesis necessary to explain what takes place in what we commonly call the external world." [16].

2.2 Abduction

Shortly after his work on the incapacities of man, Peirce introduced the concept of abduction, as a more formal term for what he previously called "hypothesis." Roughly, *abduction* occurs when we observe the properties of something, and then infer that it is a member of a class of things that share those properties.[1] Inference of class membership entails the expectations of other properties, and those expectations can be tested via future interaction. Abduction is distinguished from deduction and from induction: *Deduction* occurs when, knowing an object is a member of a class, and knowing that members of that class are characterized by a general property, we infer a property of the individual object, knowing that said property is characteristic of a general class of objects and that the individual is a member of a that class. *Induction* occurs when, knowing a property of an object, and knowing the object is a member of a particular class, we infer that property is characteristic of the general class (See Fig. 1).

The abductions Peirce most admired were those generated by the great eighteenth century chemists. Those inferences were intellectually daring leaps to multifaceted expectations regarding the results of yet-to-be-performed experiments, which were then tested in the many laboratories of a dynamic research community, and ultimately confirmed. While Peirce offered many definitions for abduction over the years, sometimes emphasizing its speculative nature, other times its logical precision, he always emphasized its crucial role in scientific discovery.

[1] See [1], and several chapters in this volume.

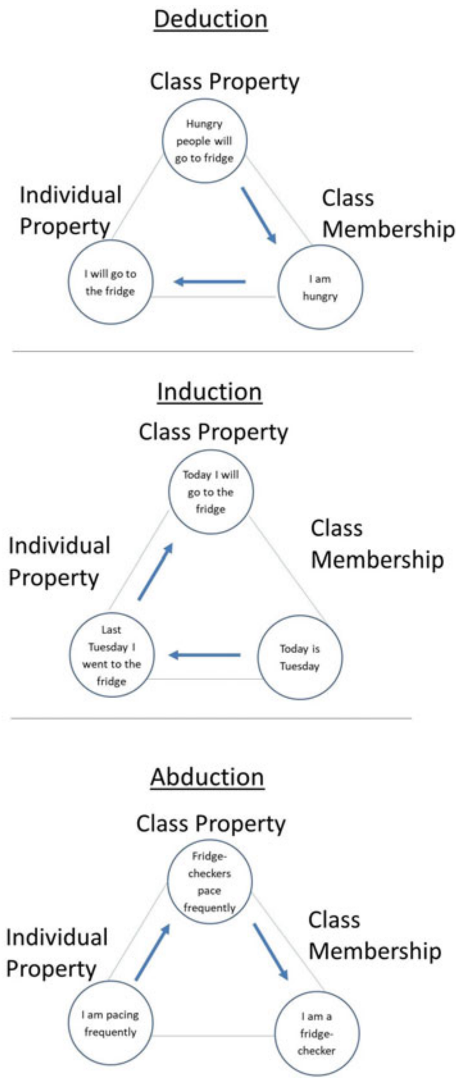


Fig. 1 The inference triangle. Given any two corners of the triangle as premises and the third corner as a conclusion, one can determine the type of inference that is occurring. From [18, 23]

Some abductions are better than others, in the sense that pursing them honestly would lead us more rapidly towards that belief that will survive long-term scientific scrutiny. Other abductions lead us only slowly towards stable belief or lead us altogether astray. We suspect that much of the confusion over what, exactly, Peirce meant by abduction is due to his frequent failure to distinguish between whether he was talking about abduction-in-general or high-quality-skillful-abduction in particular.

2.3 Belief

As a part of his assault on Descartes, Peirce discusses the nature of belief and doubt. Descartes tries to convince his readers that he can *know* what he doubts and what he believes, by virtue of a non-inferential introspective intuition. He further tries to convince his readers that those beliefs and doubts are on the other side of the unpassable chasm that separates his mind from the material world, with additional unpassable chasms between that world and the minds of others. We have already seen that Peirce rejected the existence of intuitive introspection. After rejecting the existence of intuitive introspection, Peirce anticipates the Pragmatic Maxim by asking us to consider what the consequences are of having a "belief" or a "doubt", Peirce concludes to have a "belief" is to have something *upon which we are willing to act*, whereas a "doubt" is uncertainty regarding how to act [17].

Being willing to act upon something can have very broad or very narrow implications, depending on the belief in question. And yet, regarding many beliefs we may use a key action as shorthand. A hungry man who believes his desired food is in the cupboard, will go to the cupboard and open it, continuing his behavior smoothly if the food is revealed, and acting surprised if it is not. The belief that the desired food is in the refrigerator would be similar, but with a course of action oriented towards the refrigerator instead of the cupboard. A man who has doubts about which of the two locations contains the food, in contrast, will be hesitant in taking either actions. At the extreme he will resemble Hadrian's Ass: If the hungry man has absolute doubt about the location of the food, he will not be able to take any action at all, as there is nothing upon which he is willing to act.[2]

2.4 The Pragmatic Maxim

A year after his discussion about the effects of belief and doubt, Peirce formally introduced the Pragmatic Maxim as a way to determine the meaning of *any* concept [19], "*Consider what effects, that might conceivably have practicial bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object*." Peirce's use of the word Pragmatic and "practical" has led some to conclude that Pragmatism was a seat-of-the pants, short-sighted, anti-theoretical philosophy which identified truth with short-term efficacy or even social power.[3] On the contrary, Peirce's philosophy sought to develop an enduring, coherent and ethical understanding of the world, i.e., that understanding

---

[2]American Philosophy scholars might well point out that we are slightly muddling how Peirce might be expected to present these ideas with how James might be expected to present them. As we are psychologists, building towards insights about the field of psychology, we will own any such criticisms for the time being.

[3]Peirce deplored any casuistic understanding of Pragmatism so much that he eventually felt the need to distinguish his approach by adding an extra syllable and calling it Pragmat*ic*ism. In the same spirit, N.S.T believes we could bring clarity by editing The Maxim to add a syllable to the

most likely to survive the vicissitudes of future discovery. Peirce is squarely focused on rigorous thinking and precise action: “There is no distinction of meaning so fine, as to consist in anything but a possible difference of practice.” Further, “…if one can define accurately all the conceivable experimental phenomena which the affirmation or denial of a concept could imply, one will have therein a complete definition of the concept, and *there is absolutely nothing more in it*” (excerpted from [10]).

Peirce offers a crucial corollary to his maxim: Whenever it seems as if two ideas are at play, and those two ideas are found to entail all the same conceivable effects, then only one idea is at play. In that context, let us imagine that one chemist asserted that a particular compound “contained iron” while another asserted disagreement by claiming the compound was a “ferrous composite.” If we set out to determine whether or not the two chemists *really* disagree, we would ask them about how they expect a material within their specified category to respond under various conditions: What would we find if we exposed it to a magnet? What if we immersed in pure oxygen gas? What if we put it in a spectrometer? Etc., etc., etc. If the two chemists agree regarding *all* the expected consequences-of-investigation, then regardless of how vehemently they claim to disagree, we (as third parties) will conclude that they their conceptions are the same and their disagreement illusory.

When trying to determine if there are two concepts at play or just one, we will not give any weight to ways in which the chemists might treat the substances differently that do not involve interaction with the substance itself. For example: We will *not* ask what letters they would write on a vial containing the substance, because whether one chemist would write “iron” on the vial and the other chemist “ferrous” is irrelevant. While the chemists might think of the written letter as relating to who is right and who is wrong, it is not *an effect* in the sense the pragmatic maxim; we are only interested in what would be found when we interact with (act upon) the compound itself.

2.5 Summary

At the end of his three papers on the “Incapacities”, Peirce leads us to the conclusion that self-knowledge results from inferences of the same kind as those that produce knowledge about others. However, he does not tell us what type of inference or inferences those might be, and at the time he had not even introduced the term “abduction.” He then provides us with the pragmatic maxim as a method for determining the meaning of our concepts, and provides an example of what is found when that maxim is applied to the concept of “belief.” We are now prepared to circle back

word “practic*i*al.” That would stress the fact that Peirce was using “pragmatic” in the sense of a practice, especially adherence to longstanding scientific practices of laboratory science, interlaboratory debate and cooperation, and theoretical analysis. The Pragmaticist Maxim would then read, “Consider what effects, that might conceivably have practicial bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object.”

to ask how we can understand self-knowledge within a system that has dispensed with dualistic inclinations.

3 Mental Terms in Light of Pragmatism

Treating "Know Thyself" as a directive implies that knowing-about-yourself is not the result of an innate and infallible ability, but rather is a skill to be developed over time. Such a skill would not have been conceived if we could not tell, to some extent, that some people know themselves better than others, and that some people improve the accuracy of their self-knowledge over time. How can it be that some people are better at knowing about themselves than others are? How can it be that some people get better at knowing about themselves over time? While it is hard to say *exactly* how those differences are possible, we can nonetheless work out some conditions that must be true for such differences to be possible. If knowing yourself *is* a skill one can improve, then there must be facts-of-the-matter against which self-claims push, facts which sometimes push back. Were it impossible to be wrong when inferring about yourself, or were it impossible to detect when you had been wrong, then improvement at the task would also be impossible. Invoking the pragmatic maxim: *If* "I am mad" had all the same (experimentally-investigable) effects as "I am not mad", *then* "mad" and "not-mad" would have the same meaning.

Phrased the other way around: For it to be possible to find "I am mad" true and "I am not mad" false, or vice versa, the two conceptions must have different practical consequences; it must be possible to arrange conditions to reveal effects that are entailed in one of the concepts, but not entailed in the other.

The central questions regarding self-knowledge are therefore: *By what process of inference do we arrive at self-knowledge? And, what are the effects about which we are inferring?*

Answering those questions presents many challenges, because our standard terminology is laden with dualism. Normally we would label the process of gaining self-knowledge as "introspection." But the term "introspection" implies that there is an "interior" that I am "seeing into." Of course, we do not mean literally looking into ourselves, as when examining a flesh wound. We are also *not* referring to indirect perceptions of our insides, as one might get from an x-ray or a brain-scanning device; "know thyself" discussions date back thousands of years, and are not attempts to infer about structures of brain tissue, patterns of oxygen consumption, or the like. We need a new starting place.

3.1 Revisiting "Belief"

Because Peirce's early ideas about "belief" were well developed, we can find a foundation by starting with self-knowledge questions such as "What do I believe?" or "Do I actually believe X?" Given Peirce's definition of "belief," asking such questions is an attempt to infer the nature of our habits: Upon what would I unhesitantly act? Were the right type of situation to arise, would I, in fact, unhesitantly act as if X was true? Am I *that type* of person?

This method is analogous to how a chemist, trying to infer about the *type* of substance in a bottle, is inferring about the effects that would be witnessed if certain procedures were to be performed on the chemical. "What does it look like?" It is a silvery metallic. "Would it change shape if I swirl the bottle?" It does. "Would it be attracted to a magnet?" It isn't. From that you might infer it is mercury, because mercury is a *type* of chemical for which those results would be expected. And, due to the pragmatic maxim, we know that making such an inference (concluding that the substance "is mercury") is nothing beyond inferring about the outcomes of numerous other operations yet to be performed on the substance. The chemist does not "look into" the chemical any more than the introspecting person "looks into" herself, and yet both draw meaningful conclusions about the type of thing they are dealing with.

Depending on the belief in question, our introspection may be asking about a narrow range of behavior-circumstance pairings (habits), or a nigh infinite number. A narrow example, so narrow as to seem trivial, could occur if we are being overly-introspective when seeking out food: "Do I *believe* the food is in the fridge?", we ask ourselves. "I think that I do, but I recognize that I will not know for certain until I see if the fridge is where I end up when looking for food!" The triviality makes that example seem odd. Imagining the results of such introspection being overturned in the course of events would be hard, but not impossible. However, even sticking to similar narrowly-specifiable cases, we can easily find more interesting examples. For example, we might introspect regarding political candidates, asking questions such as, "Which candidate in the election do I believe is the best, and how much doubt do I harbor?" In that example, the actual act of voting serves as a clear "moment of truth." Likely most people, especially as an election approaches, are quite firm and accurate in their answers to such questions, but people are not perfect predictors. Sometimes a voter's support for a particular candidate is not clear until they are in the voting booth and suddenly, unexpectedly, hesitate. A similar moment of truth occurs when a person wonders if he or she is "ready to leave home and go off to college." Some young people correctly infer that it will be hard for them; they struggle to leave. Others correctly infer that they are ready; they stride out confidently out, their readiness visible to all. Others firmly believe they are ready, but then hesitate, learning the truth about themselves in that moment. Others expect great difficulty, but learn the truth about themselves when departure suddenly seems easy.

For more complex beliefs, many habits may be in play. "Do I believe that it is important to help the poor when I am able?" No crucial moment verifies such a belief. You may have chance encounters with poor people in countless situations, and seek them out proactively in countless others. In each encounter several behaviors may qualify as “helping.” Also, the degree to which you are “able” might vary depending on circumstances outside the encounter in question. Do I donate to charities that help the poor? Do I assist individual poor people, either spontaneously, such as giving money to a random panhandler, or systematically, such as giving to particular poor people with whom I have social relationships? Do I encourage others to assist the poor? And, crucially, what desirable activities do I pass up in order to help the poor? The latter is crucial, as “it is important to” indicates a prioritization of helping the poor over engaging in other, less-important activities. Assuming I am engaged in an honest attempt at self-knowledge, I can consider past instances in which I acted out the behavior-vs-circumstance pattern that is “helping the poor”, as well as instances in which I did not act out that pattern. Weighing such evidence, I can infer what behaviors I am likely to exhibit in future relevant circumstances.

3.2 More Complex “Mental States”

The same basic reasoning that applies to inferences about belief must also apply to inferences about more complex mental phenomenon. Let us take a person, for example, participating in a leadership training program. The program offers training in a range of leadership skills, which our participant can choose from. One of the options is “Communication”, which prompts the self-directed question “Am I a good communicator?” We have written extensively about what “communication” means to field researcher in animal behavior [4, 22], and offer a definition of communication that answers the question: What would I need to see (and under what conditions would I need to see it) to confirm that I am a good communicator? It is, in other words, what a clear-thinking person (in Peirce’s sense) would mean when using the term.

“Communication” is when one organism has a behavioral system (a system that matches its behavior to its circumstances), and a second organism has a system that interacts with the first organism’s system to facilitate the first system’s function. For example: One organism, perhaps a type of monkey, might have an “anti-predator” behavioral system that responds to seeing a snake by trying to getting off the ground, and responds to seeing a hawk by trying to be close to a tree trunk. A second monkey might have a “signaling” behavioral system that responds to seeing a snake by producing a certain call and responds to seeing a hawk by producing a different call. If the first monkey responds to the “snake call” as it would respond to seeing a snake, and responds to the “hawk call” as if it had seen a hawk, then we are dealing with a communicatory system. Crucially, everything about those two systems, and everything about their interaction can be observed under naturally occuring or experimentally arrangable circumstances. Or, phrased to match our earlier discussion of the Pragmatic Maxim: That analysis provides all the conceivable experimental phenomena which “the monkeys communicate” could imply, and *there is absolutely nothing more in it.*

Generalizing back to humans: To be an *effective communicator* is to be the type of person who reliably makes those around them better match their behavior to their circumstances better, by virtue of interacting with those other peoples' behavioral systems. Someone trying to determining whether or not to focus on improving "communication" skills, must infer whether or not she is the type of person who is already adept at assisting others in that way, or if she needs to improve her ability to help others in that way.

Certainly, in trying to determine what type of person she is, our prospective leader faces practical limits in determining the truth of the matter; she cannot now, nor in any future, observe herself under all possible circumstances entailed in such a complex concept. However, no deeper limitations prevent her from gaining such knowledge. There is no other thing, no ethereal "internal soul" about which she is attempting to infer.

Note that, while we will never experience every possible circumstances in which our communicatory ability could be tested, given enough time we will have opportunity to observe ourselves across a large swath of varied circumstances, and determine whether our hypothesis proves out. Just as with other meaningful claims about the world, it is possible—when making mental claims—to investigate honestly and determine when you are wrong.

This "introspective" inference is, in form, identical to what we would do for a colleague who was looking for advice about whether or not to take the "communication" class. When trying to infer "What type of person is *he*?" we would need to think about the extent to which we have seen that person exhibit parts of the larger pattern that is "communication." We ask ourselves, "Have we seen him successfully vary behavior this way and that, depending on the circumstances, so as to help others better match their world? Have we seen him fail to do so?" And from the resulting evidence, we infer will happen in future interactions. Concluding that he *is* a good communicator is not a definitive judgement about past interactions, or about some nebulous brain state, or about a quality of his intangible soul, but rather a prediction regarding what will happen in future interactions.

4 Self-Knowledge and Types of Inference

What type of inference is entailed in the above examples? In the simple examples, we can contort the situation to make it *seem* like any of the three types of syllogistic inference resulted in the relevant conclusion. For example, we might imagine using deduction, induction, or abduction to answer the question of whether or not I will go to the fridge to look for cake. If I know that hungry people always look in the fridge (premise 1, class property), and that I am hungry (premise 2, class membership), then I can infer via *deduction* that I will look in the fridge (inference of individual property). If I know that today is Tuesday (premise 1, class membership), and that last Tuesdays I looked in the fridge (premise 2, individual property), I can infer via *induction* that Tuesdays are days when I will look in the fridge (inference of class property). If I know my stomach is growling (premise 1, individual property), and that fridge checkers tend to have growling stomachs (premise 2, class property), then I could infer via *abduction* that I am a fridge-checker (inference of class membership) (See Fig. 2). While the conclusion of each of those inferences clearly differs if ones looks closely, at a rougher level they all seem to suggest I will check the fridge.

The more complex the mental state we are concluding about, the harder to imagine all types of inference being leading me to the same conclusion about myself. That is because, as the mental phenomenon in question gets more complex, we can no longer act like we are merely concluding about whether we will engage in a single crucial

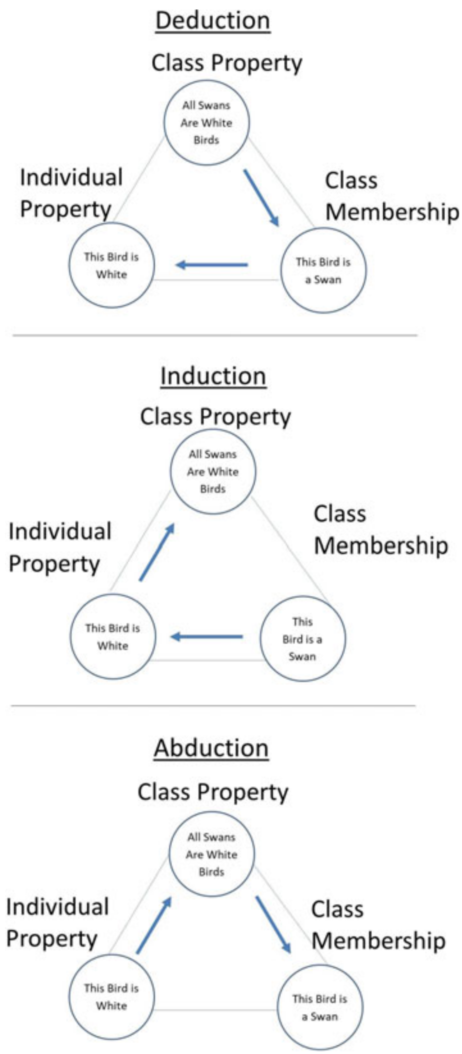


Fig. 2 The inference triangle, using example from text (modified from Thompson [23]). Note that while all three inferences may initially seem to involve the same conclusion, upon closer inspection the conclusion is quite different depending on the mode of inference. In the first case I am concluding something about myself as an individual, in the second case I am concluding something about a property of a group, and in the third case I am concluding that I am a member of a type

action; the more complex the mental phenomenon in question, the more clearly we are trying to infer about a broad pattern of action across numerous circumstances.

Do I have a good memory? That is a question about the matching of numerous behaviors to past circumstances. Am I brave? Again, while some moments might be crucial in making the judgement, the question is about far more than a single act. The directive to "know thyself" is not directing you to know the exact way you will act in one particular situation, it is directing you to know what *type* of person you are. To draw conclusions about what type of person you are, based upon your having traits characteristic of members of that type, is, by definition, abduction.

Even acknowledging that in vivo inference is typically of a complex variety, as inferential processes are rarely encountered in pure isolation (see Shook, this volume), it is clear that the key act of gaining insight about our nature results from an abductive process.

4.1 False Self-Belief

Just as we generally understand the risks of error entailed in inferring about the minds of others (trying to classify what type of people *they* are based on having observed some amount of *their* behavior), we should generally acknowledge all the same risks when we infer about ourselves. Both are abduction. The person wondering if he believes in helping poor people might not have accurate ideas about his past behavior in regards to helping poor people; he might not be honest with himself about what the pattern "helping poor people" looks like; and even without those challenges he might come to an inaccurate conclusion about future behavior, because abductive inference is fallible. Similarly, when trying to infer about others you would encounter the same limitations: How accurate is your knowledge of the other person's past behavior? How accurate is your understanding of the thing you are concluding about? How accurate an inference can you draw from whatever base of ideas you start with?

Abduction-concerning-the-self carries all the same risks as abduction about the mental states of others, because those risks are inherent in any abduction. Stated simply: We can be mistaken regarding each premise, and we can be mistaken about the strength of the inference drawn from the premises. The same fundamental problem is faced by our archetypal abductor, the bench chemist, in his efforts to infer the nature of a substance: She may not have correctly read her instruments, or may have otherwise gained inaccurate ideas about how the substance responded to past tests (premise 1, individual property). She may be inaccurate regarding the "properties" of different known substances (premise 2, type property). And even if she was accurate regarding experimental results *and* was accurate regarding the properties of known chemicals, she may still draw an incorrect conclusion (regarding type membership). And what would the consequences be of an incorrect abduction, i.e., of concluding incorrectly regarding the identity of a chemical? Abduction to an incorrect conclusion entails failing to correctly anticipate the result of future experiments involving the chemical. Abduction is a probabilistic form of inference, and the fundamental risks faced by people engaged in abductive inference remain the same across subject matter domains.

5 Can Others Know Us Better Than We Know Ourselves?

If self-knowledge is achieved through everyday powers of inference, why have arguments-for-a-special-power-of-introspection have been so convincing in the past. Presumably, one important reason past generations have believed in special powers of introspection is that many people feel like they know themselves quite well. That feeling must be accounted for.

First, we must acknowledge the extensive evidence in the experimental-psychology literature that people are not particularly good at predicting how they will act in novel situations (extensively summarized in [12]). Those data suggest that people feel good about introspection because of hindsight and confirmation bias, or related phenomena. When people evaluate their own introspective abilities, they may inaccurately remember past anticipatory inferences, with a bias towards thinking they correctly anticipated their behavior even when they did not (hindsight bias). Also, they might correctly remember past anticipatory inferences, but do so selectively, attending primarily to instances in which they anticipated correctly, and neglecting instances in which their anticipations were incorrect (confirmation bias). Those biases can easily combine with other biases, such as the Dunning-Kruger effect, whereby people with marginal competence at a task dramatically overestimate their competence. Thus, it is possible people believe they have excellent self-knowledge simply because they overestimate their own ability.

However, even after acknowledging that people are not, on average, as good at self-knowing as they seem to think, they might be even worse at knowing about others. That is, some people might feel like they are very good at anticipating what they will do in various situations, because they are so much worse at anticipating what others will do, and they are thinking in relative terms. Are there reasons why abductions about ourselves *might* be more reliable, on average, than abductions about others? At least two reasons seem obvious: First, we might have more data upon which to base self-related inferences and, second, we might have access to more useful information upon which to base our self-related inferences.

5.1 Access to More Observations

Regarding the first possibility, people have a wealth of experience being around themselves. Presumably we have each been around ourselves more often than we have been around any other individual. To the extent that can accurately draw upon that past experience, we may be able to reach more reliable conclusions than we could reach about others.

If this explanation was true, then the seeming advantage of introspection would be an extreme example of the principle that we are generally better at anticipating the behavior of those we have known a long time, and worse at anticipating the behavior of those we have known only a short time. This explanation also matches our intuition that people with memory problems, either temporary or permanent, should show deficits in self-knowledge.

5.2 Access to More Useful Observations

Regarding the second possibility, we note that we often do not present ourselves honestly to those around us. If there are behaviors, or inclinations towards behaviors, that we hide from those around us (e.g., situations in which we lie to others about our past behavior, or situations in which we make special effort to not be observed), then other people, making an inference based on the best information they have available, could be at a disadvantage. This possibility isn't about the amount of data available for our inferences, but about access to particularly useful data.

The problem with this second line of inquiry is that even if we are typically able to accurately recount our past behavior, nothing guarantees such efforts are always accurate. If Freud was right about one thing, it was that people sometimes lie to themselves, deploying various strategies designed to avoid coming to terms with *exactly* the types of self-knowledge insights we are discussing. Whether or not Freud was right in his description of how so-called "defense mechanisms" operate, he was correct that we frequently do not know why we respond in certain ways in certain situations. If Freud was right about a second thing, it was that a properly trained professional may gain insights into the motivations of others, and thereby learn exactly the types of things that "intuitive introspection" is supposed to reveal. The professional psychoanalyst tries to gather a lot of data about the subject's past behavior and circumstances, and deploys strategies designed to reveal examples the client would normally not reveal to others. These can including examples the client would not normally choose to face on his or her own. The practice is based on the assumption that, provided sufficient time, and the right conditions, a stranger may become more proficient at anticipating your behavior than you are.

5.3 Summary

Once we have rejected the notion of an infallible introspective intuition, we are left with the conclusion that even if people are quite good at introspection, those insights are still best understood as the result of an inferential process. We cannot take refuge in the idea that introspection works well because it provides privileged access to uniquely insightful information. Nothing guarantees that the crucial information is always available to introspection, and nothing guarantees that others cannot have that information, given the proper circumstances. Abduction about the self has the same basic limitations as abduction about others.

6 Conclusion

We have argued that inference of the "know thyself" variety—so-called "introspective" inference, about who you are as a person—takes the form of abduction. Such inferences do not differ in kind from the abductive inferences by which we draw conclusions about the minds of others. In both cases, we are ultimately inferring what type of person someone is—we are inferring membership in a general class—which leads to expectations of what that person would do under a variety of different conditions. Whether inferring about ourselves or about others, were we to follow up the inference in the manner of honest investigators, concerned with whether or not our ideas were *true*, we would proceed to create various situations and observe reactions. In both cases we can be mistaken in well understood ways.

The pragmatic maxim dictates that the *meaning* of a concept is nothing more than all conceivable practical effects the concept entails. If we are correct that the process of testing claims regarding mental terms entail tests of behavior under various circumstances, then that is the entire meaning of the concepts. Understanding self-knowledge and other-knowledge as abductive inference helps us see the ubiquity of abduction in everyday life. As highly social animals, we are navigating a constant maze of probable inferences regarding ourselves and those around us, with some of those inferences being verified, others being rejected, and still others remaining untested because the crucial circumstances never arise.

Also, by understanding that both self-knowledge and other-knowledge results from abductive inference we can better understand how to use mental terms within a *science* of psychology. We cannot give into the crude operationalism of some behaviorists on the one side, nor to the pervasive crypto-dualism of cognitive psychologists on the other side [7, 11]. We must understand our mental terms as references to complex clusters of potential experimental results, while not allowing into our thinking anything that is inherently non-investigable. Only then can we create a solid foundation for the field.

References

1. Campos, D.G.: On the distinction between Peirce's abduction and Lipton's Inference to the best explanation. Synthese **180**, 419–442 (2011)
2. Charles, E.P.: ToM 101: a guide to mainstream theory of mind research [Review of the book Theory of mind: How children understand others' thoughts and feelings, by M. J. Doherty]. PsycCRITIQUES **54**(15) (2009)
3. Charles, E.P.: Myths and missteps in "theory of mind" dogma [Review of the book Against theory of mind, by I. Leudar & A. Costall, Eds.]. PsycCRITIQUES **55**(25) (2010). https://doi.org/10.1037/a0020186
4. Charles, E.P.: Seeing minds in behavior: descriptive mentalism. Rev. Gen. Psychol. **15**(3), 267–276 (2011)
5. Charles, E.P., Bybee, M.D., Thompson, N.S.: A behaviorist account of emotions and feelings: making sense of James D. Laird's Feelings: The perception of self. Behav. Philos. **39**, 1–16 (2011)

6. Charles, E., Dege, M.: The beyond must also be useful: the burden of alternative approaches. Integr. Psychol. Behav. Sci. **42**(2), 194–199 (2008)
7. Charles, E.P., Golonka, S., Wilson, A.D.: The most important thing neuropragmatism can do: providing an alternative to 'cognitive' neuroscience. In: Shook, J.R., Solymosi, T. (eds.) Pragmatist Neurophilosophy: American Philosophy and the Brain, pp. 127–149 (2014)
8. de Waal, C. (ed.): American New Realism 1910–1920. Thoemmes Press, Bristol, United Kingdom (2005)
9. Doherty, M.J.: Theory of Mind: How Children Understand Others' Thoughts and Feelings. Psychology Press, New York (2009)
10. Houser, N.: Peirce's Post-Jamesian Pragmatism. Eur. J. Pragmatism Am. Philos. [Online], III-1 (2011). URL: http://journals.openedition.org/ejpap/866
11. Heft, H.: Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James' Radical Empiricism. Erlbaum Inc, Mahwah, NJ (2001)
12. Laird, J.D.: Feelings: The Perception of Self. Oxford University Press, New York (2007)
13. Leudar, I., Costal, A.: Against Theory of Mind. Palgrave Macmillan, New York (2009)
14. Miller, S.A.: Theory of Mind: Beyond the Preschool Years. Psychology Press, New York (2012)
15. Peirce, C.S.: Questions concerning certain faculties claimed for man. J. Speculative Philos. **2**, 103–114 (1868)
16. Peirce, C.S.: Some consequences of four incapacities. J. Speculative Philos. **2**, 140–157 (1868)
17. Peirce, C.S.: Grounds of validity of the laws of logic: further consequences of four incapacities. J. Speculative Philos. **2**, 193–208 (1869)
18. Peirce, C.S.: The fixation of belief. Popular Sci. Monthly **12**, 1–15 (1877)
19. Peirce, C.S.: How to make our ideas clear. Popular Sci. Monthly **12**, 286–302 (1878)
20. Smith, L.D.: Behaviorism and Logical Positivism: A Reassessment of the Alliance. Stanford University Press, Stanford, CA (1986)
21. Thompson, N.S.: The many perils of ejective anthropomorphism. Behav. Philos. **22**, 59–70 (1994)
22. Thompson, N.S.: Communication and natural design. In: Thompson, N.S. (series ed.), Owings, D., Beecher, M. (vol. eds.) Perspectives in Ethology, vol. 12: Communication. Plenum, New York (1997)
23. Thompson, N.S.: Signs and designs. Behav. Philos. **46**, 25–46 (2018)
24. Thompson, N.S., Charles, E.P.: Interview with an old new realist. In: Charles, E.P. (ed.) A New Look at New Realism: The Psychology and Philosophy of E. B. Holt, pp. xxxi–lviii. Transaction Publishers, Piscataway, NJ (2011)

The Role of Collections of Objects in Abduction

Patricia Turrisi

Abstract The abductive stage of logical investigation benefits from direct contact with objects of thought, especially material objects in aggregates. The ubiquity of collecting activity and collections of non-utilitarian material objects in ancient as well as contemporary settings, with increasingly deliberate attention to material objects as implements of thought, demonstrates these benefits. This essay focuses on the logica utens of collecting and the role of collections in detection and discovery.

Keywords Logical investigation · Inquiry · Logica utens · Logica docens · Charles Peirce · Signs · Representations · Art of memory · Recollection · Memory · Mary Carruthers · Paleozoic era · Non-utilitarian objects · Objective idealism · Museums · Collections · Collecting · Niles Eldredge · Hoarding behavior · Process philosophy · Research · Discovery · Scott Herring · Cicero · Theory of the universe · Unifying theory of nature · Charles Darwin · Alfred Lord Wallace

Abductions may certainly be formulated in the inner space of the mind, in an empty room in the absence of objects of perception and thought, as when one has a "bright idea" inspired by a dream or spontaneous intuition. Such a process substitutes mental representations for objects and relies on meta-data and its selective representation of the features of objects, focusing more typically on already well-formed conceptions, while neglecting unmediated experiences. Objects of thought themselves are generally richer in content than our selective perceptions and memories make them out to be. Rather than presuming a sterile environment for thought that only includes precise representations in the form of propositions that may serve as the premises and conclusions of abductive arguments, let us imagine a lavish environment in which a plenitude of possible objects of thought resides. At times, the economy of research has presided too much over the generation of abductions. Economy in research eliminates not only lines of inquiry but also access to inquiry. Let us instead imagine that economic restrictions on time and resources will take care of themselves or that

P. Turrisi (✉)
University of North Carolina, Wilmington, USA
e-mail: turrisip@uncw.edu

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_13

there will be breakthroughs in methods of research that permit broader and deeper inquiries than are possible now. In doing so, we may turn to the issue of how an abundance of samples and specimens foster the generation of abductions.

In his 1898 Cambridge Conference Lectures, Charles Peirce asserted that the first rule of reason is, "Do not block the way of inquiry" [1, p. 178]. In her article, "Peirce's First Rule of Reason and the Process of Learning," Amy McLaughlin supplements the interpretation of Peirce's rule made by Susan Haack which focuses on "the (ideal) product of (completed) inquiry, truth" [2, p. 233]. McLaughlin expands Haack's interpretation by first noting Douglas Browning's understanding of Peirce as a "philosopher of process," the key feature of process philosophy being that "the universe is essentially to be understood as creative, organic, and temporal." McLaughlin stresses that Peirce "describes learning as a 'gain of experience'," motivated by dissatisfaction with one's own experience, understood as "a hunger, or thirst" more accurately than as a mere absence of opinion [3, pp. 233–234]. In surveying the kinds of motivations and ideals attributed to collectors, I took this point of view to heart. Collections and collecting are neither random nor aberrant human behavior. Rather, as we say about aspects of certain systems, they are a feature, not a bug, of inquiries seeking sustenance for hunger or thirst "to find out, from the consideration of what [one] already know[s], something else which [one does] not know" [4, p. 2].

The Peirce papers at Houghton Library consist of more than 100,000 pages in various partial stages of cataloguing and preservation. Peirce studied chemistry, worked as an assistant in the Harvard Observatory for four years and for the United States Coast and Geodetic Survey for thirty years. He amassed collections of observations and measurements as a researcher in his work in both these institutions. His scientific and philosophical papers display a vast accumulation of facts and observations from which he drew hypotheses as well as new methodologies for investigating scientific and philosophical questions. His self-described method of work began, in each case of a given question coming up for discussion, with "writing a collation upon it," including arguments for and against one and then the other side of the matter [5, p. 386]. His method then called for him to "draw up a list of the general methods in which a solution of the problem might be sought." In other words, he collected every thought and perspective on the problem at hand and kept it before his eyes. He further processed these lists, referring to each conception of the problem as "a seedling tree, which must have water, nutriment, sunlight, shade, and air and frequent breaking of the ground about it, in order that it may grow into something worthy of respect," also calling each iteration a "digestion" [5, p. 387]. It seems clear that Peirce took pains to allow plentiful reflection. While the resources for undergoing any given investigation might be limited, he saw no reason to censure his ambitions or expectations for nurturing ideas.

1 *Logica Utens*: Logical Doctrine Without the Systems and Traditions

Aside from studies of Peirce, logic in this century exists primarily in the realm of meta-data, in text and discourse (propositions) as a phenomenon exclusive of reality, especially materiality. Logic appears to be a separate field from ontology (especially materiality as a subcategory of ontology) though there may be overlapping concerns. However, these concerns characterize logic as formal and, for the most part, as a set of purely intellectual operations that omit any reference to content. For example, Thomas Hofweber claims, in an essay intended to be generic and universal, that,

Overall, we can thus distinguish four notions of logic:

- (L1) the study of artificial formal languages
- (L2) the study of formally valid inferences and logical consequence
- (L3) the study of logical truths [more or less acceptable rules of logical operations]
- (L4) the study of the general features, or form, of judgements [6]

Such definitions of the realm of logic confine it to a content-free set of formal rules and operations. An understanding of Peirce's logic of abduction, which is admittedly only nominally formal but rich in suggestions regarding content, is best constructed not through the four ideas of logic listed above, but rather from Peirce's own writings and the relevant literature in Peirce studies. This essay distinguishes, as Peirce does, between *logica docens* and *logica utens*. *Logica docens* is the systematic scientific study of logic; *logica utens* is the rational process that is antecedent to any systematic study of the forms of logic. The acts of reasoning I will cite in this essay, apart from Peirce's, are generally examples of *logica utens*. *Logica utens* does not lack a logical doctrine or logical tenets. *Logica utens* is merely reasoning without a theory of itself.

Though its users do not deliberately employ formal logical systems, they do make deliberate and self-controlled inferences whose forms "[they] approve as calculated to lead to the truth" [7].

Anyone can engage in abduction without knowing the theory of it. Anyone can adopt a conclusion provisional to its eventual proof. An abduction "is the process of forming an explanatory hypothesis." Peirce identifies an abduction as "the only logical operation which introduces any new idea" [1]. Abductive inference has this form:

> The surprising fact, C, is observed;
>
> But if A were true, C would be a matter of course, Hence, there is reason to suspect A is true. [8]

Logica utens is likely to focus on the surprising fact of C initially. Why or how is C possible? How did C get there? Why does C appear to have the form that it does? One makes guesses that explain C. My motive for discussing collections is that collections are reasonable to study as the source of A, as the examples illustrate.

2 Objects Speak to Us: Peirce's "One Intelligible Theory of the Universe"

The "one intelligible theory of the universe," Peirce claims, is objective idealism, which holds that "matter is effete mind, inveterate habits becoming physical laws." Rejecting both the autonomy of matter from mind and the autonomy of mind from matter, he underlined their relationship as one that evolves from irregularity to regularity. He claims that

> The old dualistic notion of mind and matter, so prominent in Cartesianism, as two radically different kinds of substance, will hardly find defenders today. Rejecting this, we are driven to some form of hylopathy, otherwise called monism. Then the question arises whether physical laws on the one hand and the psychical law on the other are to be taken
>
> (A) as independent, a doctrine often called monism, but which I would name *neutralism*; or,
>
> (B) the psychical law as derived and special, the physical law alone as primordial, which is *materialism*; or,
>
> (C) the physical law as derived and special, the psychical law alone as primordial, which is *idealism.*
>
> The materialistic doctrine seems to me quite as repugnant to scientific logic as to common sense; since it requires us to suppose that a certain kind of mechanism will feel, which would be a hypothesis absolutely irreducible to reason, – an ultimate, inexplicable regularity; while the only possible justification of any theory is that it should make things clear and reasonable.
>
> Neutralism is sufficiently condemned by the logical maxim known as Ockham's razor, i.e., that not more independent elements are to be supposed than necessary. By placing the inward and outward aspects of substance on a par, it seems to render both primordial.
>
> The one intelligible theory of the universe is that of objective idealism, that matter is effete mind, inveterate habits becoming physical laws. [9]

On the way to acquiring inveterate habits, is it not then possible that matter expresses itself in signs that are *in*-exact, that is, neither measurable nor expressible in mathematical signs as one might expect in the "exact sciences"? For example, since the ostensibly regular phenomenon of the speed of an object is often computed using the formula: speed = distance/time, any number of objects in motion can be partially characterized with reference to one of the regularized habits of matter. However, countless other effects that are materially expressed do so more or less irregularly, that is, in patterns that are more or less habitual and, therefore, only more or less measurable. Positivist philosophers have taught us that entities that are not measurable may as well not exist for the discipline of philosophy, but *logica utens* practitioners live in a world *filled* with such objects. Their effects appear in perceptual judgments that are difficult to analyze precisely through symbolic systems. Nonetheless, they guide our understanding of the world and especially our situational awareness of our own personal worlds. Sometimes, they can be characterized in general terms, through the language of nonverbal communication as studied in proxemics, kinesics and other fields such as anthropology and even the neurosciences. I am proposing that not only does nonverbal communication take place between sentient beings such as ourselves, but also between sentient entities such as human beings and material objects, especially those that have only partially developed "inveterate habits." After all, the ability to calculate the speed of an object only reveals a partial truth about it. There is much more to learn. Thus, in the spirit of Peirce's first rule of reason, let us consider a variety of objects of thought that are accessible to us through our interactions with collections—quantifiable as well as unquantifiable, material as well as nonmaterial.

3 Objects Speak to Us: Hominin Foundations of the Urge to Collect

Non-verbal capacities to express responses to the environment that appear in early (and contemporary) humans predate the evolution of human symbolic communication systems.

Neurobiological studies locate the source of nonverbal responses in the limbic system, a part of the brain thought to be the oldest in evolutionary time. Researched by "anthropologists, archaeologists, biologists, linguists, psychiatrists, psychologists, semioticians, and others who have studied human communication from a scientific point of view," *The Nonverbal Dictionary of Gestures, Signs & Body Language Cues* is a rich assembly of terms that elaborate the conditions under which human beings discern signs that are durable, material, aromatic, auditory, tactile, vestibular, visual and edible [10]. The arrangements, combinations, contrasts, rhythms, and sequences of signs have their own systems. The perception and use of space, and the system of spatial signs, speak to us, albeit without words and in terms difficult to measure. The *Nonverbal Dictionary* distinguishes between efferent and afferent signs, acknowledging in many entries both signs that humans send nonverbally and signs that are sent to humans nonverbally, by other human beings *and objects alike.* Afferent and efferent signs are part of a closed nervous system loop that consists of nerves that accompany sensations, decisions and reactions. The nervous system registers afferent signs as "incoming" sensations while efferent signs are characterized as "outgoing" decisions and reactions, mostly expressed in muscle movement and glandular activity, distinct nerve fibers registering affect and effect in this system. As observers of responses to material objects that give pleasure, for example, the *affect* is a "subjective feeling" of pleasure (as in the notion, "mmm, this smells good"), while the *effect* may be a change in hormonal levels, or other facial or bodily expression or action, with reference to the source of pleasure. Charles Peirce would have been excited to study the relationships that nonverbal specialists now understand have evolved between objects and perceivers; for example, in the case of "new car smell":

> *Aroma cue.* A scented consumer product designed to mimic the leather, rubber, plastic, and vinyl aromas of a show-room-new motor vehicle interior.
>
> *Usage.* We find the synthetic odor of new car smell pleasant because it contains chemical analogs of natural plant resins, animal esters, and sexual steroids.
>
> *Evolution.* New car smell, which may be sprayed from aerosol cans, was developed by International Flavors and Fragrances of New York, which supplies odor cues for Downey Fabric Softener and Colgate's Irish Spring soap. [11]

The "new car smell" experience is based on an entirely manufactured, chemically precise substance that becomes a sign when as-signed to a new car and is understood by a perceiver to stand for a set of affairs, the distinctive aroma of cars in a showroom or on a test drive. Pleasure is somewhat measurable in brain activity, but to understand the *experience* of pleasure it provides (relating to emotion and mood, which are afferent sensations) requires recognition of a set of *relations* between the consumer, the chemistry of "new car smell," the limbic system of the consumer, the adaptive value of proximity to plant resins, animal esters and sexual steroids in human existence, and the material object that prompts an almost instantaneous suggestion of the meaning of these relations, the "new car" itself. The perceptual judgments that arise from experiencing "new car smell" are not merely isolated events of odor detection, but also a *guess* as to the meaning of a complex network of relationships.

The urge to interact with material objects is itself a nonverbal behavior as are the acts of collecting and maintaining collections. The *Nonverbal Dictionary* has adopted the term *object fancy* to indicate:

> 1. The desire to pick up, handle, and hold a material object, especially a consumer product of elegant design. 2. The urge to touch, own, arrange, collect, display, or talk about a manufactured human artifact. 3. The motivation for compulsive shopping. *Usage*: Products "speak" to us nonverbally as tangible, material gestures. Their *design features* (e.g., the *shine*, *shape*, and *smoothness* of a platinum bracelet) send compelling messages to capture our attention. That we respond to their appeal shows in the sheer number of artifacts we possess. Our personality may be caricatured by the object(s) we desire, e.g., jewelry, boats, shoes, and so on. We may hold treasured artifacts with two hands, in a gentle, caressing embrace between the tactile pads of our thumbs and forefingers. Forever beckoning from TV monitors, mail order catalogues, and shelves, products gesture until we answer their call. [12]

"Object fancy" does not describe a utilitarian desire to fulfil one's basic needs for food, shelter or defense, though it may well be an adaptive desire that promotes well-being from other perspectives. Standard timelines of human development cite tool use and cooking as the first major breakthroughs in human cultural evolution. Early humans in East Africa as early as 2.6 million years ago used hammer stones to produce sharp flakes with which to process new foods. However, a number of non-utilitarian lithic objects have come to light in various parts of the world from previously excavated collections held in museums [13]. These objects were collected as part of past archaeological explorations but have only recently been curated as significant objects in themselves. Against denials of the symbolic value and, indeed, the emergence of a symbolic understanding of objects previous to *homo sapiens,* or more than 75,000 years old, Moncel et al. claim, in their research report, "Non-Utilitarian Lithic Objects from the European Paleolithic," that:

> The odd and non-utilitarian lithic objects occurring in the Paleolithic [~ 2.6 million – 12,000 years ago] sites indicate that they retained the attention of hominins, and were treated with special care for some time, at least from the time that they were carried from the natural source to the living site. Because these objects are related to a range of activities apart from subsistence, they are significant for our understanding of the day-to- day life of the Paleolithic people.
>
> What the motivation was for collecting extraordinary items is impossible to assess in early hominins, and even in modern humans. Picking up an extraordinary stone and bringing it back "home" may represent a casual action, or alternatively, this action may be related to various individual or collective concerns like play, aesthetic feeling, symbolic communication or magical-religious practice, among others. [13, p. 25]

These non-utilitarian, "extraordinary" objects include fossils, quartz crystals and other minerals, stones with cupules, objects with anthropomorphic shapes, pebbles, amber, black stones, steatite, smooth pebbles, shells and shell fragments. Collection activity in early hominins is difficult to characterize as unequivocally indicative of specific cognitive capacities, yet the authors conclude that:

> The frequency of these objects increases with time [from the Lower to the Upper Paleolithic Ages] and, although their presence dramatically increases after the appearance of Homo sapiens, they are still apparent in early hominins and Neanderthal sites….
>
> With the early hominins this is indicated simply by the collection of odd, unusual, extraordinary objects, picked up at the same time as the necessary goods for survival. Such behavior has not been observed so far among other primates, even though they use wood or stone tools and play with pieces of these common raw materials (Goodall 1989; Joulian 2005; Morgan and Abwe, 2006; Mercador et al. 2007). [13, p. 34]

Early collection activity "was mainly centered around objects distinguishable by their color, transparency, shape, surface texture, etc." while, after the appearance of *Homo sapiens*, "the collection of unusual objects, like fossils, fragments of rocks with odd shades or morphology, [and] small colored pebbles continued." *Homo sapiens* also began to transform unusual raw materials "in a more systematic manner" for symbolic or aesthetic purposes. Their apparently great interest in exceptional materials "led to increasingly distant searches, involving movement of objects and/or people," with the selection of objects based on color as before, but also increasingly "on the possible effects of transparency or luster, as well as on the degree of softness" [13, p. 34].

4 Systems of Collection: The Sheet of Assertion

In this section and the next, I compare one of Peirce's techniques for organizing objects of thought with a pre-modern system of re-collection, i.e., the art of memory, both of which acknowledge the purpose of collections as means of generating new thought.

An abduction terminates in an assertion that is a supposition about the relationships between perceptions, or collections of perceptions, or of regularities as regards other regularities, or the relationship between something in evidence and its cause or explanation. The Sheet of Assertion is a device Peirce imagined in order to construct a mode of expressing sets of assertions for further examination:

> 1903 | A Syllabus of Certain Topics of Logic | Peirce, 1903, p. 15; CP 4.397
>
> It is agreed that a certain sheet, or blackboard, shall, under the name of *The Sheet of Assertion*, be considered as representing the universe of discourse, and as asserting whatever is taken for granted between the graphist and the interpreter to be true of that universe. The sheet of assertion is, therefore, a graph.
>
> 1903 [c.] | Logical Tracts. No. 2. On Existential Graphs, Euler's Diagrams, and Logical Algebra | CP 4.430
>
> What we have to do […] is to form a perfectly consistent method of expressing any assertion diagrammatically. The diagram must then evidently be something that we can see and contemplate. Now what we see appears spread out as upon a sheet. Consequently our diagram must be drawn upon a sheet. We must appropriate a sheet to the purpose, and the diagram drawn or written on the sheet is to express an assertion. We can, then, approximately call this sheet our *sheet of assertion*.
>
> 1903 [c.] | Logical Tracts. No. 2. On Existential Graphs, Euler's Diagrams, and Logical Algebra | CP 4.432
>
> A certain sheet, called the *sheet of assertion*, is appropriated to the drawing upon it of such graphs that whatever may be at any time drawn upon it, called the *entire graph*, shall be regarded as expressing an assertion by an imaginary person, called the *graphist*, concerning a universe, perfectly definite and entirely determinate, but the arbitrary creation of an imaginary mind, called the *grapheus*. [14]

Let us imagine the shelves in the storeroom of a museum as a proto-sheet of assertion, with its specimens systematically withdrawn from their natural state by collectors and curators.

These objects propel us toward much more than propositions about individual specimens, though such assertions may be added to the sheet in the form of labels, indexes, catalogues and oral reports. The organization of the collection of objects permits references by each individual in the set to others with which it is categorized, and each subset of collections to other subsets of collections. Furthermore, the system by which collectors have organized each collection and sub- collection may be conjectured as another feature of the collections. Collections additionally tell us something about their collectors, which may be interesting from historical and cultural perspectives. The minds that categorize objects are worth understanding because they add to the community of scholarship that attempts to render the world intelligible. Now imagine other kinds of collections such as the approximately 2.5 billion square feet of self-storage rentable space in the USA, the approximately 55,000 museums in the world, 75,000 antique stores on the internet, 25 million sellers on eBay, or 291 million monthly active Pinterest users. Add to these your own library, your bookmarks on your browser, or your collection of anything you please. You might be the subject of object fancy, as primitive as picking up pretty rocks as you walk from one place to another, or you might collect things that may be useful in making sense of some question or problem. In the case of the sheet of assertion, Peirce re-collects that which the graphist and the interpreter agree is the widest Universe of Reality. Moreover, "the diagram must then evidently be something that we can see and contemplate." Items on the sheet can be added or erased, put into different configurations, and added to again and again as an inquiry progresses. For this reason, it is useful to introduce the method and purpose of the pre-modern art of memory, which was intended to create a structure "in trained memory that built character, judgment, citizenship, and piety" [15, p. 11]. Just as Peirce's technique proposes a collection-preserving artifact that serves as a prompt for thoughtful inquiry, the art of memory gives a method for collecting the means for a thoughtful human disposition.

5 Systems of Collection: The Art of Memory and Memory as Re-collection

The art of memory as practiced in pre-modern times through the Middle Ages bears little resemblance in method or purpose to memory schemes that are everywhere present in the modern world. "Brain games," for example, claim to aid memory by providing stimulation that prevents brain atrophy. "Mental athletes" practice mnemonics to improve memory while everyone else relies on external devices because they no longer have to store "information," that is, facts or bits and pieces of data. The meanings and uses of memorized "information" are supposed to be unnecessary or irrelevant to the practice of memory. A typical description claims that,

> Mnemonics was a once common art of using particular techniques to improve the strength of your memories. These techniques were popular back when people really needed to rely on their mind's ability to keep track of information. [16]

However, the purpose of the classical art of memory was neither as utilitarian nor as trivial as its contemporary practice imagines. In the Greek tradition of mnemonics, perceptions, regardless of how they were presented to the mind, were "encoded as *phantasmata,* 'representations' or a 'kind of *eikôn*'" [15, p. 20]. The goal of recollection was to invent and compose in the present, not to reproduce a record of past events or to "keep track of information." The further goal of such inventions and compositions was to enrich and enhance the human condition through enrichment by communities of educated scholars ("textual communities") whose universe of discourse was a shared grasp of key texts. Books, language and other human artifacts themselves were cues to "existent truths," not fundamentalist tracts with no further need of interpretation or application [15, p. 11]. Furthermore, writing was not considered a supplement (or replacement) of memory.

Rather, as Cicero wrote:

> Memory … is in a manner the twin sister of written speech [*litteratura*] and is completely similar to it [*persimilis*], [though] in a dissimilar medium. For just as script consists of marks indicating letters and of the material on which those marks are imprinted, so the structure

of memory, like a wax tablet, employs places [*loci*] and in these gathers together [*collocat*] images like letters. [15, p. 18]

To be clear, writing, whether on an actual or imagined surface, was understood to be distinct from "writing down." Mary Carruthers states, in *The Book of Memory*, that "the ability to write is not always the same thing as the ability to compose and comprehend in a fully textual way, for indeed one who writes (a scribe) may simply be a skilled practitioner, employed in a capacity to that of a professional typist today" [15, pp. 17–18]. The purpose of *memoria* was to stimulate and facilitate thinking about moral character and civic responsibility in pre-modern times as now, an activity that required creative interpretation and not merely a mechanical adherence to formulae:

> *Memoria*, as these writers understood and practiced it, was a part of *litteratura*: indeed it was what literature, in a fundamental sense, was for. Memory is one of the five divisions of ancient and medieval rhetoric; it was regarded, moreover, by more than one writer on the subject as the "noblest" of all these, the basis for the rest. *Memoria* was also an integral part of the virtue of prudence, that which makes moral judgment possible.
>
> Training the memory was much more than a matter of providing oneself with the means to compose and converse intelligently when books were not readily to hand, for it was in trained memory that one built character, judgment, citizenship, and piety. [15, p. 11]

Carruthers clarifies the "distinction between memory understood as the ability to reproduce something exactly ("rote") and memory as recollection or reminiscence" [15, p. 22]. Regardless of whether the scheme of retrieval was natural or artificial, "both sorts of schemes are heuristic, inventive and investigatory in the classical sense" (of a learning and making practice) [15, p. 24]. She emphasizes that the representations or images stored in memory each have a "cognitive function," that is, its user understands each to represent a certain thing. Peirce defines a sign in a comparable manner. Moreover, his method of work may be understood as a kind of re-collection in Carruthers' terms. Citing a predominant model of memory work involving inscription on a block of wax, as discussed in Plato's *Theaetetus* and in Aristotle's *Metaphysics*, Carruthers notes that the root word Plato uses for "seal," (*sême*) or sign, is "a mark by which something is known." Carruthers explains that these copies, whether stored in the mind or stored on a wax tablet, are "sort-of" images, not precisely replicas of the somethings that are known, but evocative rather, a "quasi-picture" representation that calls something to mind, perhaps the initial perception, but more useful for invention, composing and learning, a trove of associations contextualized by the process of memorial collection in the first instance [15, p. 27]. The pre- modern memory artist's skill was most evident in the weaving together (*textus*) of *phantasma* into original discoveries such as Thomas Aquinas did in his compilations (collections) of Patristic texts made for Pope Urban or the sudden breakthrough in his argument against the Manichees in *Summa theologica*.

These remarks suggest that in any medium that memorializes thought, there are *collections* of almost-representations that more or less cohere with one another. There is much food for thought here as we consider books, libraries, and other collections of manuscripts and addenda curated (in material and mental forms) by textual communities. However, it may be a challenge to understand how one ever determines what to collect, which somethings to focus upon when "we may wish to remember something we see or hear or conceive in our own minds."[1] Two examples should illustrate both that such choices are useful means of generating abductions and that a choice to collect an item is itself abductive, the abduction in general being that it will be useful in generating abductions.

6 Scientific and Artistic Collections. Niles Eldredge: Paleontologist and Cornet Collector

Niles Eldredge has been Curator in the Division of Paleontology at the American Museum of Natural History since 1969. In 1972, he and Stephen Jay Gould formulated the theory of Punctuated Equilibrium in distinction from Charles Darwin's gradualist theory of evolution. Eldredge also studies the relationships between global extinctions in the geological past and the current crises in biodiversity [17]. His initial work with Gould took flight from his specialization in mid-Paleozoic phacopid trilobites. Trilobites are an extinct class of Paleozoic marine arthropods consisting of over 25,000 described species [18]. The American Museum of Natural History owns specimens of trilobites from the Cambrian through Permian Periods (roughly 542–251 million years ago), from regions as diverse as North America, Africa, Europe, Asia, Australia, Russia, South America and China (roughly everywhere). More recently, Eldredge has been comparing patterns and processes in biological evolution with a similar process in material cultural evolution using his collection of over five hundred antique cornets. He told reporter Margaret Wertheim, "I'm doing now with cornets exactly what I used to do with trilobites: measuring, analyzing and cataloging the myriad gradations of their forms" [19]. He further commented that, as his collection grew, he "began to realize that 'the basic cornet designs are essentially packages of information, of cultural material information, similar to species in the biological realm." He wondered if patterns similar to those found in the fossil record could be seen in the evolution of cultural material artifacts. His work on cornets identified a set of characteristics that he has perceived as capturing "the essential anatomical features of the cornet" in order to further explore "the way new ideas and designs spread through an instrument population." Ilya Temkin, Eldredge's student and an expert on the psaltery, an instrument played in Latvia, Lithuania, Estonia, Finland and Russia, is applying methods of classification used to study pearl oysters to search for patterns in psaltery evolution. Both scientists have derived intriguing hypotheses concerning differences between biological evolution and the evolution

[1] In Theaetetus, when Socrates introduces the imaginary block of wax in our minds, he states, "Let us call it the gift of the Muses' mother, Memory, and say that *whenever we wish to remember something we see or hear or conceive in our own minds*, we hold this wax under the perception of ideas and imprint them on it as we might stamp the impression of a seal ring. Whatever is so imprinted we remember and know so long as the image remains; whatever is rubbed out or has not succeeded in leaving an impression we have forgotten and do not know" [2, p. 24].

of the manufacture of musical instruments. One example is lateral transmission, which is common among bacteria and plants but otherwise rare among other organisms. Lateral transmission of characteristics of musical instruments occurs when one manufacturer creates an innovative feature, others copy it, and it spreads throughout a region. Eldredge and Temkin identify patterns in manufacture and pair them with economic and social patterns in the period of manufacture to discover more about these patterns in order to promote further guesses about how these patterns emerged.

7 Scientific and Artistic Collections: Museums

In his book, *Museums: A History*, John E. Simmons traces the history of collections from hoards and artifacts in tombs to modern museums in order to "discern worldviews and, in some instances, the motivations of collectors" [20, p. xiii].

Simmons cites Alma S. Wittlin's argument that "an account of premuseum collections based on chronology or geography was not as useful as categorizing collections by function because 'collections both reflect and affect human ways of life'" [20, p. 11]. Wittlin proposed six categories, one of which was "collections intended to stimulate curiosity and inquiry." It would be impractical to survey every collection in every museum in the world, but it may suffice to review several examples of museum collections intended to stimulate such curiosity and inquiry. The first is the Alexandrian Temple of the Muses, which "formed the conceptual basis of museum development beginning in the Renaissance" in its use of objects as sources of knowledge [20, p. 31]. Reliable historical evidence of the contents and layout of the Alexandrian Temple of the Muses does not exist, but imaginative and conjectural accounts describe it as based on shrines to the Muses that promoted learning and civic harmony. Simmons claims it is "best described as a combination of a school of scholars, a research institution and a library" [20, p. 34]. Alexander sent specimens from Persia and India to his tutor, Aristotle, including many described by Aristotle in *Historia Animalium* and from which Aristotle developed a taxonomic system of some 450 animals he arranged in terms of their relative perfection on the *scala naturae* [20, p. 35]. Historians have determined that the Alexandrian Temple of the Muses contained, among other accommodations for scholars, an observatory, a zoo, and a botanical and meditation garden. The library contained a number of scrolls. Ptolemy II "was reputed to have required travelers arriving in Alexandria to surrender any manuscripts in their possession so that they could be copied for the library" [20, p. 37]. The manuscripts collected in the library were edited and corrected to produce the most authentic versions of the Greek literature of classical antiquity, including stories, songs, poems and plays that represented a rich array of understandings of human life, relationships between humans and gods, and stories of war, travel and exploration [20, p. 37].

With exploration and the technology that allowed better preservation of specimens and artifacts, museums in the modern period became places where collections were established for research. Major modern museums display only a small portion of the objects they own for study.

For example, as of 2015, the Victoria and Albert Museum displayed 233,742 objects and works of art but had 2,044,441 books, drawings, prints, photographs and archives available for study [21]. As of 2013, the Field Museum in Chicago contained 25 million artifacts but displays less than one per cent at any given time. According to the Field's Associate Curator of Birds, "the rest are one of the best research collections in the world for all types of biological and cultural diversity" [22]. Many donations to museums have their origin in private collections originally amassed for research purposes. Charles Darwin kept a heated greenhouse where he grew plants for his experiments. Preserved bird specimens collected by Darwin on his voyage on the HMS Beagle were first given to the Zoological Society of London and later distributed to the British Museum and other institutions. Two hundred of his bird specimens now reside in the UK's Natural History Museum and seven other collections. Half of his original collections are unaccounted for, but it is certain that, over his lifetime, he compared myriad specimens that led him to formulate theories about the origins of differences between varieties of organisms and their relationships to geographical and environmental conditions [23]. Darwin's co-presenter of the theory of natural selection, Alfred Russel Wallace, likewise collected natural history specimens, especially plants, insects, beetles and birds, many of which are now held in several dozen institutions internationally. The significance of evolutionary interpretations of specimen collections in originating, supporting and amplifying what has become the unifying theory of the life sciences cannot be overstated. Moreover, as Steven Lubar claims, in *Inside the Lost Museum*, it is inadvisable merely to replace the objects and specimens in museum collections with digital copies because new inquiries can arise from old objects. Arguing for the preservation of large collections, he identifies Harvard University's Museum of Comparative Zoology's collection of mice, mouse collections at the Smithsonian Institution, the Museum of Vertebrate Zoology in Berkeley and others whose specimens are over a century old. Asking "but what use are they today?" he notes the response by Hopi Hoekstra, Curator of Mammals at the Museum of Comparative Zoology: "the mice they collected are invaluable, documenting 'more than a century of dynamic relationships between deer mice and their environment'," allowing scientists to study population genetics, and "to answer questions about climate change and how mice respond to changing environments" [3, p. 274]. More significantly, such collections allow scientists to *ask* questions about the relationship between the past, present and future.

8 Collecting as Disorder/Disorder as "Useful Destabilization"

Should anyone believe that collectors are all of a piece and mentally ill rather than in the main quite serious about the uses of collections for sparking inquiries about the nature of things, it is important to clear the air: close to a century of ideology about "hoarding" and "hoarders" is rooted in indefensible claims.

The DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders*) in 2013 formally recognized hoarding behavior as "Hoarding Disorder," which it characterized as "persistent difficulty discarding or parting with possessions, regardless of their actual value, as a result of a strong perceived need to save the items and to distress associated with discarding them" [24, p. 235]. This authoritative manual distinguishes what it terms a hoarding disorder from normal collecting when hoarding results in clutter in active living areas "such that their intended use is substantially compromised." An extreme form of hoarding disorder "consists of excessive collecting, buying or stealing of items that are not needed or for which there is no available space" [24, p. 235]. The pathology of the supposed disorder assumes it has a neurobiological origin despite a complete lack of evidence. In "Collyer Curiosa: A Brief History of Hoarding" Scott Herring refutes the interpretation of one supposed case of a hoarding disorder, Collyer Brothers Syndrome, as a representation of an "anal-erotic" trait or as biochemical imbalance. He instead "treats hoarding as a chapter in the unfinished cultural history of disorder and 'gross disorganization'" [25, pp. 161–2]. He offers a plausible alternative hypothesis to regarding the condition variously labelled as collector's mania, pathological collecting and, especially, chronic disorganization, as a mental disorder [25, pp. 159–160]. Rather, he demonstrates how a pernicious interpretation of Harlem and of migrant neighborhoods gave rise to a global interpretation of the dubious category of excessive collecting as a dangerous mental aberration.[2]

Compare the "hoarding disorder" diagnosis of a collector whose possessions "are not needed" with a museologist's justification of the method of using objects as evidence:

> The object as keeper of the truths of the past, as the most truthful historian, appeals to the museum curator. The curator can reveal those truths and make the objects speak. [3, p. 280]

Lubar describes perceiving old objects in new contexts, the abduction of the significance of objects as they occur alongside others:

> Most curators combine methods [of analysis, classification and description] with a visceral understanding of artifacts. Nicholas Thomas of the Cambridge Museum of Archaeology and Anthropology describes the museum method as more about "discovery" than critical enquiry: "it often involves finding things that were not lost, identifying things that were known to others, or disclosing what was hidden or repressed.
>
> The encounter with arrays of objects is usefully destabilizing." [3, p. 280]

[2]See also Martin F. Manlansan IV, "The 'Stuff' of Archives: Mess, Migration, and Queer Lives" in Radical History Review Issue 120 (Fall 2014), 94–107 for an ethnographic account of undocumented Asian and Latino queer immigrants living in Jackson Heights, Queens. Manlansan derives an account of transitory lives that are organized disorganizations of "the tragic, the repulsive, the uncomfortable, the banal, and the seemingly trashy, off-kilter and bodily practices that may hint at political potentials, gesture to alternative narrative, and enable an openness to multiple futures." ([26], p. 106).

In Peirce's terms, such destabilizing encounters unblock the way of inquiry. Lubar explains that storerooms are places where curators practice "museum sense" in its purest form, which they also call "object-feel" or "a good eye." In storerooms, curators may experience "object-love" (akin to "object fancy"): "a curators' passion for his or her collection, and for every object in it." Lubar notes, "Curators love their collections, and storerooms are 'shaped by the emotion attached to the objects they house'" [3, p. 101]. While it would be a mistake to conflate museum storerooms with dwellings full of seemingly useless leavings, the value of hundreds of artifacts in storage areas may be inestimable for research.

References

1. Peirce, C.S.: Abduction. Harvard lectures on pragmatism: lecture VI. MS [R] 314. In: Bergman, M., Paavola S. (eds.) The Commens Dictionary: Peirce's Terms in His Own Words. New Edition (1903). http://www.commens.org/dictionary/term/abduction. Accessed 23 May 2019
2. McLaughlin, A.: Peirce's First Rule of Reason and the Process of Learning. In: Thellefsen, T., Sorensen, B. (eds.) Peirce in His Own Words: 100 Years of Semiotics, Communication and Cognition, pp. 229–234. De Gruyter Mouton, Boston (2014)
3. Lubar, S.: Inside the Lost Museum: Curating Past and Present. Harvard University, Cambridge (2017)
4. Peirce, C.S.: The Fixation of Belief. Popular Sci. Monthly **12**, 1–15 (1877)
5. Turrisi, P.: Peirce's Method of work. In: Thellefsen, T., Sorensen, B. (eds.) Peirce in His Own Words: 100 Years of Semiotics, Communication and Cognition, pp. 385–391. De Gruyter Mouton, Boston (2014)
6. Hofweber, T.: Logic and Ontology. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2018/entries/logic-ontology/. Accessed 18 Aug 2018
7. Peirce, C.S.: Reason's Rules. MS [R] 596. In: Bergman, M., Paavola S. (eds.) The Commens Dictionary: Peirce's Terms in His Own Words. New Edition (1902–03 [c.]). http://www.commens.org/dictionary/term/logica-utens. Accessed 23 May 2019
8. Peirce, C.S. Abduction. Harvard lectures on pragmatism: lecture VI. MS [R] 315. In: Bergman, M., Paavola S. (eds.) The Commens Dictionary: Peirce's Terms in His Own Words. New Edition (1903). Accessed 23 May 2019
9. Peirce, C.S.: Objective idealism. In Bergman, M., Paavola S. (eds.) The Commens Dictionary: Peirce's Terms in His Own Words. New Edition. http://www.commens.org/dictionary/term/objective-idealism. Accessed 23 May 2019
10. Givens, D.: The Nonverbal Dictionary of Gestures, Signs & Body Language Cues. http://center-for-nonverbal-studies.org/htdocs/6101.html. Accessed 31 Aug 2019
11. Givens, D.: New Car Smell. The Nonverbal Dictionary of Gestures, Signs & Body Language Cues. http://center-for-nonverbal-studies.org/htdocs/newcar.htm. Accessed 31 Aug 2019
12. Givens, D.: Object Fancy. The Nonverbal Dictionary of Gestures, Signs & Body Language Cues. http://www.nonverbal-dictionary.org/2013/01/object-fancy.html. Accessed 31 Aug 2019
13. Moncel, M.-H., Chiotti, L., Gaillard, C., Onoratini, G., Pleurdeau, D.: Non-utilitarian lithic objects from the European Paleolithic. Archaeol Ethnology Anthropol Eurasia **40**(1), 24–40 (2012)
14. Peirce, C.S.: Sheet of Assertion. In: Bergman, M., Paavola S. (eds.) The Commens Dictionary: Peirce's Terms in His Own Words. New Edition. http://www.commens.org/dictionary/term/sheet-of-assertion. Accessed 16 May 2019
15. Carruthers, M.: The Book of Memory: A Study of Memory in Medieval Culture, 2nd edn. Cambridge University, Cambridge UK (2008)

16. Handel, S.: Mnemonics: The Forgotten Art of Memory. https://www.theemotionmachine.com/mnemonics-the-forgotten-art-of-memory/. Accessed 31 Aug 2019
17. Eldredge, N.: American Museum of Natural History, New York. https://www.amnh.org/research/staff-directory/niles-eldredge. Accessed 31 Aug 2019
18. Gallery of Trilobites. American Museum of Natural History, New York. https://www.amnh.org/research/paleontology/collections/fossil-invertebrate-collection/trilobite-website/gallery-of-trilobites. Accessed 31 Aug 2019
19. Wertheim, M.: Scientist at Work—Niles Eldredge; Bursts of Cornets and Evolution Bring Harmony to Night and Day March 9, 2004, Section F, Page 3 New York Times
20. Simmons, J.E.: Museums: A History. Rowman & Littlefield, Lanham, Maryland (2016)
21. Size of the V&A Collections. Victoria and Albert Museum. http://www.vam.ac.uk/content/articles/s/size-of-the-v-and-a-collections/. Accessed 31 Aug 2019
22. Klocksin, K.: How Much of the Field Museum's Collection Is Actually On Display? WBEZ91.5 Chicago (2013). https://www.wbez.org/shows/curious-city/how-much-of-the-field-museums-collection-is-actually-on-display/485cb7d3-b46c-434b-ae84-8f6ae8ad808f. Accessed 31 Aug 2019
23. Voyage of HMS Beagle (1831–1836) Natural History Museum. https://www.nhm.ac.uk/our-science/collections/zoology-collections/bird-skin-collections/bird-skin-collection-hms-beagle.html. Accessed 31 Aug 2019
24. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 5th edn. American Psychiatric Association, Arlington, VA (2013)
25. Herring, S.: Collyer Curiosa: a brief history of hoarding. Criticism **53**(2), 159–188 (2011)
26. Manlansan IV, M.F.: The 'Stuff' of archives: mess, migration, and queer Lives. Radical Hist. Rev. **120**, 94–107 (2014)

Our Themes on Abduction in Human Reasoning: A Synopsis

Emmanuelle-Anna Dietz Saldanha, Steffen Hölldobler, and Luís Moniz Pereira

Abstract Psychological experiments have shown that humans do not reason according to classical logic. Therefore, we might argue that logic-based approaches in general are not suitable for modeling human reasoning. Yet, we take a different view and are convinced that logic can help us as an underlying formalization of a cognitive theory, but claim rather that classical logic is not adequate for this purpose. In this chapter we investigate abduction and its link to human reasoning. In particular we discuss three different variations we have explored and show how they can be adequately modeled within a novel computational and integrated, cognitive theory, the Weak Completion Semantics.

1 Introduction

Originally, one of the objectives of using logic in artificial intelligence and knowledge representation and reasoning was the formalization of human and commonsense reasoning. During the past decades the original objective shifted out of focus. The problem description was not specified adequately, possibly because there was little or no communication with cognitive scientists. Non-classical logic approaches exist,

---

The authors are mentioned in alphabetical order.

E.-A. Dietz Saldanha (✉) · S. Hölldobler
Faculty of Computer Science, TU Dresden, 01062 Dresden, Germany
e-mail: sh43@posteo.de

S. Hölldobler
e-mail: sh@iccl.tu-dresden.de

S. Hölldobler
North-Caucasus Federal University, Stavropol, Russian Federation

L. M. Pereira
NOVA Laboratory for Computer Science and Informatics, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
e-mail: lmp@fct.unl.pt

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_14

however most of them are purely theoretical and not applied to actual human case studies. Instead, artificial examples have been constructed, which only show that a theory works within some very specific context. But what is the value of a cognitive theory that has never been evaluated against the conclusions that humans actually draw?

Taking this observation as a starting point, several human reasoning episodes, ranging from the suppression and the selection task [8, 9], to spatial, syllogistic [5] and counterfactual reasoning [6, 7] have been successfully modeled under a new, computational and integrated, cognitive theory, the *Weak Completion Semantics*. In a nutshell, under the *Weak Completion Semantics* the following steps are taken for a given human reasoning scenario:

1. Reasoning towards a (logic) program representation.
2. Weakly completing the program.
3. Computing its least model under Łukasiewicz logic.
4. Reasoning with respect to the least model.
5. If necessary, applying abduction.

The *Weak Completion Semantics* is based on ideas first presented in [22, 23]. Unfortunately, these first ideas contained a technical bug, which was corrected in [13] by switching from three-valued Kleene logic [16] to three-valued Łukasiewicz logic [18]. As shown in [13], each weakly completed program admits a least model which can be computed as the least fixed point of an appropriate semantic operator. Unsurprisingly, it turned out that some human reasoning tasks require abduction and, hence, the *Weak Completion Semantics* has been extended with abduction.

In this chapter, we will avoid introducing formal definitions, and assume the reader to be familiar with logic and logic programming. The interested reader is referred to [12, 17]. The goal of this chapter is rather to give the intuitions behind the different facets that abduction can take within human reasoning. By means of four episodes of human reasoning, *the suppression task, counterfactual reasoning, contextual reasoning* and *reasoning with obligation and factual conditionals*, we will show that abductive reasoning is required. Interestingly, a different variation of abduction applies for each of the four cases in order to be adequately modeled under the *Weak Completion Semantics*.

2 The Suppression Task

The *suppression task* is a famous psychological experiment originally carried out by Byrne [2] showing that humans suppress previously drawn conclusion if new information becomes available.

Suppose that participants are told that

if she has an essay to finish then she will study late in the library and that

she will study late in the library.

Then, the subjects are asked whether they are willing to conclude that *she has an essay to write.* As Byrne reports, 71% of the subjects are willing to draw this conclusion.

Following [22, 23], the conditional is translated into the program

$$\{library \leftarrow essay \wedge \neg ab_1,\ ab_1 \leftarrow \bot\},$$

whose weak completion is[1]

$$\{library \leftrightarrow essay \wedge \neg ab_1,\ ab_1 \leftrightarrow \bot\},$$

where *essay* denotes that *she has an essay to finish* and *library* that *she will study late in the library*. ab_1 is an *abnormality* predicate that can deal with possible exceptions to the rule, and is introduced as needed in the program representation, depending on the conditional.

The weakly completed program admits a least model under (three-valued) Łukasiewicz logic,[2] where *library* and essay are mapped to unknown and ab_1 is mapped to false. Adding the fact $library \leftarrow \top$ to the program changes the least model in that library is now mapped to true, but essay is still unknown. Hence, we should not add it but rather consider *library* as an observation that needs to be supported by an explanation. In the given context, the minimal explanation being $\{essay \leftarrow \top\}$. If added to the program, we obtain the weakly completed program

$$\{library \leftrightarrow essay \wedge \neg ab_1,\ essay \leftrightarrow \top,\ ab_1 \leftrightarrow \bot\},$$

whose least model maps *library* and *essay* to true and ab_1 to false. Reasoning with respect to this model allows to abductively conclude *essay*, which most subjects did.

Hypothesis 1 Humans reason abductively.

In another group of the same experiment participants were told:

If she has an essay to finish, then she will study late in the library.
If she has a textbook to read, then she will study late in the library.

All other information being equal to the previous experiment, the number of subjects willing to conclude that *she has an essay to write* drops to 13% [2]. The previously drawn conclusion is suppressed as soon as the second conditional is given. Modelling this reasoning episode similarly to before, we obtain the weakly completed program

---

[1] One should observe that under completion as defined in [4], $essay \leftrightarrow \bot$ would be an element of this set. The *Weak Completion Semantics* however does not map undefined relations to false but rather considers them unknown.

[2] In the sequel, all models are computed with respect to three-valued Łukasiewicz logic where not explicitly mentioned otherwise.

$$\{library \leftrightarrow (essay \wedge \neg ab_1) \vee (textbook \wedge \neg ab_2),\ ab_1 \leftrightarrow \bot,\ ab_2 \leftrightarrow \bot\},$$

where *textbook* denotes that *she has a textbook to read*. Considering *library* again as an observation to be justified, there are two minimal explanations, viz. $\{essay \leftarrow \top\}$ and $\{textbook \leftarrow \top\}$. Reasoning credulously, i.e. when one explanation suffices to conclude *essay*, then one would conclude that *she has an essay to write*. However, humans do not do this. Their conclusions appear to be modelled adequately when skeptical abduction is applied.

Hypothesis 2 Humans prefer skeptical over credulous abduction.

As shown in [8, 14], all twelve cases of the suppression task can be adequately modelled under the *Weak Completion Semantics*. For six of these cases, skeptical abduction is required.

3 Counterfactual Reasoning

Consider the following scenario [1]: *President Kennedy was killed. There was a lengthy investigation about whether Oswald or somebody else shot the president. In the end, it was determined that Oswald did it.* Which of the following conditionals do humans accept easily?

If Kennedy is dead and Oswald did not shoot Kennedy then someone else did.
If Oswald had not shot Kennedy then someone else would have.

According to [1], humans accept the former, but reject the latter conditional.

In this case, the background knowledge is encoded in the program

$$\begin{array}{lll} \{\ kennedy \leftarrow oswald \wedge \neg ab_o, & ab_o \leftarrow \bot, & \\ \ \ kennedy \leftarrow someone\_else \wedge \neg ab_s, & ab_s \leftarrow \bot, & oswald \leftarrow \top\ \}, \end{array}$$

where *kennedy* denotes that *Kennedy was killed*, *oswald* denotes *Oswald shot*, *someone_else* denotes *somebody else shot*, and ab_o as well as ab_s are introduced abnormality predicates. As the weak completion of this program we obtain

$$\begin{array}{llll} \{\ kennedy \leftrightarrow (oswald \wedge \neg ab_o) \vee (someone\_else \wedge \neg ab_s), & & & \\ \ \ ab_o \leftrightarrow \bot, & ab_s \leftrightarrow \bot, & oswald \leftrightarrow \top & \}, \end{array}$$

whose least model maps *oswald* and *kennedy* to true and ab_o and ab_s to false.

The antecedent $\neg oswald$ of the second conditional is false in this least model. Hence, the conditional is a *counterfactual*. In order to evaluate this counterfactual we must revise the background knowledge. The *Weak Completion Semantics* does this in a straightforward and minimal way by replacing the positive fact $oswald \leftarrow \top$ with the negative assumption $oswald \leftarrow \bot$ in the program. Consequently, the least model of the weakly completed modified program maps *oswald*, ab_o and ab_s to false and *kennedy* to unknown. If we evaluate the second conditional with respect to this least model we find that the antecedent $\neg oswald$ of the conditional is true, whereas its consequence *kennedy* is unknown. Thus, the conditional itself is evaluated as unknown.

If we evaluate the first conditional with respect to the least model of the weak completion of the original program, then its antecedent $kennedy \wedge \neg oswald$ is also false as the literal $\neg oswald$ is mapped to false. If the original program is revised as before, then $\neg oswald$ is mapped to true, but now *kennedy* is unknown with respect to the least model of the weakly completed modified program. Hence, the antecedent $kennedy \wedge \neg oswald$ is unknown. However, *kennedy* can be explained by abduction. The only minimal explanation is $\{someone\_else \leftarrow \top\}$. Adding this explanation to the revised program and weakly completing it, we obtain

$$\begin{array}{l}\{\ kennedy \leftrightarrow (oswald \wedge \neg ab_o) \vee (someone\_else \wedge \neg ab_s),\\ \quad ab_o \leftrightarrow \bot, \quad ab_s \leftrightarrow \bot, \quad oswald \leftrightarrow \bot \quad someone\_else \leftrightarrow \top\ \},\end{array}$$

whose least model maps *someone_else* and *kennedy* to true and *oswald*, ab_o and ab_s to false. Under this model the antecedent $kennedy \wedge \neg oswald$ as well as the consequent *someone_else* of the first conditional are mapped to true. Consequently, the conditional itself is evaluated to true.

Hypothesis 3 If needed, humans minimally revise their background knowledge before applying abduction.

Various counterfactual reasoning scenarios, such as Pearl's *firing squad* scenario [19] and Byrne's *forest fire* scenario [3] can also be adequately modeled by applying *minimal revision followed by abduction* under the *Weak Completion Semantics*, as detailed in [7].

4 Contextual Reasoning

How can we prefer explanations that explain the normal cases to those explanations that explain the exceptional cases? How can we express that some explanations have to be considered only if there is some evidence for considering the exception cases? We want to avoid having to consider all explanations if there is no evidence at all for considering exception cases. On the other hand, we don't want to state that all exception cases are false.

Let us consider the famous Tweety example from [20] which is as follows:

Birds usually fly. Tweety and Jerry are birds.

According to [20], if nothing else is known about Tweety and Jerry, then we can deduce that both can fly.

This scenario can be easily modeled in the *Weak Completion Semantics*. The background knowledge is encoded in the program

$$\{fly(X) \leftarrow bird(X) \wedge \neg ab(X),\ ab(X) \leftarrow \bot,\ bird(tweety) \leftarrow \top,\ bird(jerry) \leftarrow \top\},$$

where X a variable, $fly(X)$ denotes that X *can fly*, $bird(X)$ that X *is a bird* and $ab(X)$ is an abnormality predicate. The corresponding ground program is

$$\begin{array}{llll}\{ fly(tweety) \leftarrow bird(tweety) \wedge \neg ab(tweety), & ab(tweety) \leftarrow \bot, & bird(tweety) \leftarrow \top, & \\ \ \ fly(jerry) \leftarrow bird(jerry) \wedge \neg ab(jerry), & ab(jerry) \leftarrow \bot, & bird(jerry) \leftarrow \top & \}.\end{array}$$

Weakly completing this program we obtain

$$\begin{array}{llll}\{ fly(tweety) \leftrightarrow bird(tweety) \wedge \neg ab(tweety), & ab(tweety) \leftrightarrow \bot, & bird(tweety) \leftrightarrow \top, & \\ \ \ fly(jerry) \leftrightarrow bird(jerry) \wedge \neg ab(jerry), & ab(jerry) \leftrightarrow \bot, & bird(jerry) \leftrightarrow \top & \}.\end{array}$$

whose least model maps $bird(tweety)$, $bird(jerry)$, $fly(tweety)$ and $fly(jerry)$ to true and $ab(tweety)$ as well as $ab(jerry)$ to false. Reasoning with respect to this least model we conclude that *Tweety and Jerry can fly*.

Suppose we observe

Jerry does not fly.

This observation cannot be explained in the usual abductive framework as specified in [15]. In such a framework the set of abducibles is defined to be the set of undefined relations with respect to the given program, where a relation or ground atom A is said to be *undefined* given a program if (the ground instance of) the program does not contain a rule of the form $A \leftarrow Body$, where $Body$ is a conjunction of literals. In the program above all relations ($fly(jerry)$, $fly(tweety)$, $bird(jerry)$, $bird(tweety)$, $ab(jerry)$, $ab(tweety)$) are defined.

However, under the *Weak Completion Semantics*, negative assumptions like $ab(jerry) \leftarrow \bot$ can be *defeated* by positive facts like $ab(jerry) \leftarrow \top$. The weak completion of a negative assumption $A \leftarrow \bot$ and its *defeater* $A \leftarrow \top$ is $A \leftrightarrow \top \vee \bot$, which is semantically equivalent to $A \leftrightarrow \top$. We believe that humans may defeat negative assumptions. To this end we allow defeaters of negative assumption to be abducibles in an extended abductive framework. Now, the observation *Jerry does not fly* can be explained by the minimal explanation $\{ab(jerry) \leftarrow \top\}$. In other words, *Jerry does not fly because it is an abnormal bird.*

Hypothesis 4 Humans may defeat negative assumptions.

Consider this extended scenario:

Usually birds can fly, but kiwis and penguins cannot.
Tweety and Jerry are birds.

This background information can be encoded by the program

$$\begin{array}{lll}\{ fly(X) \leftarrow bird(X) \wedge \neg ab(X), & & \\ \ \ ab(X) \leftarrow kiwi(X), & ab(X) \leftarrow penguin(X), & \\ \ \ bird(tweety) \leftarrow \top, & bird(jerry) \leftarrow \top & \}.\end{array}$$

The corresponding ground program is

$$\begin{array}{ll} \{\, \mathit{fly}(\mathit{tweety}) \leftarrow \mathit{bird}(\mathit{tweety}) \wedge \neg \mathit{ab}(\mathit{tweety}), & \\ \;\; \mathit{fly}(\mathit{jerry}) \leftarrow \mathit{bird}(\mathit{jerry}) \wedge \neg \mathit{ab}(\mathit{jerry}), & \\ \;\; \mathit{ab}(\mathit{tweety}) \leftarrow \mathit{kiwi}(\mathit{tweety}), & \mathit{ab}(\mathit{tweety}) \leftarrow \mathit{penguin}(\mathit{tweety}), \\ \;\; \mathit{ab}(\mathit{jerry}) \leftarrow \mathit{kiwi}(\mathit{jerry}), & \mathit{ab}(\mathit{jerry}) \leftarrow \mathit{penguin}(\mathit{jerry}), \\ \;\; \mathit{bird}(\mathit{tweety}) \leftarrow \top, & \mathit{bird}(\mathit{jerry}) \leftarrow \top \qquad \}. \end{array}$$

Weakly completing this program we obtain

$$\begin{array}{l} \{\, \mathit{fly}(\mathit{tweety}) \leftrightarrow \mathit{bird}(\mathit{tweety}) \wedge \neg \mathit{ab}(\mathit{tweety}), \\ \;\; \mathit{fly}(\mathit{jerry}) \leftrightarrow \mathit{bird}(\mathit{jerry}) \wedge \neg \mathit{ab}(\mathit{jerry}), \\ \;\; \mathit{ab}(\mathit{tweety}) \leftrightarrow \mathit{kiwi}(\mathit{tweety}) \vee \mathit{penguin}(\mathit{tweety}), \\ \;\; \mathit{ab}(\mathit{jerry}) \leftrightarrow \mathit{kiwi}(\mathit{jerry}) \vee \mathit{penguin}(\mathit{jerry}), \\ \;\; \mathit{bird}(\mathit{tweety}) \leftrightarrow \top, \quad \mathit{bird}(\mathit{jerry}) \leftrightarrow \top \qquad \}. \end{array}$$

whose least model maps *bird*(*jerry*) and *bird*(*tweety*) to true and all other ground atoms to unknown.

Suppose we observe

Jerry flies.

This observation can be justified by the minimal explanation

$$\{\mathit{kiwi}(\mathit{jerry}) \leftarrow \bot,\ \mathit{penguin}(\mathit{jerry}) \leftarrow \bot\}.$$

One should observe that both atoms, *kiwi*(*jerry*) and *penguin*(*jerry*), must be mapped to false in order to map *ab*(*jerry*) to false as well, which is a prerequisite for mapping *fly*(*jerry*) to true. In other words, we need to assume that *Jerry is not a kiwi and not a penguin* in order to conclude that *Jerry flies*.

There are several problems with this approach. Firstly, to the best of our knowledge there are currently 41 known classes of birds which do not fly. Secondly, only specialists in biology might know these 41 classes. Thirdly, there may be classes of flightless birds which we are unaware of. Hence, it is unlikely that humans consider all known exceptions before concluding *Jerry flies*.

In order to overcome these problems, the *Weak Completion Semantics* was extended with a new truth-functional operator *ctxt* (called *context*) in [10], whose meaning is specified in Table 1. The meaning of *ctxt* can be understood as a mapping

Table 1 The truth table for *ctxt L*, where *L* denotes a literal

L	*ctxt L*
$\top$	$\top$
$\bot$	$\bot$
U	$\bot$

from three-valuedness to two-valuedness, which allows to locally capture *negation as failure* [4] under *Weak Completion Semantics*. Using the *ctxt* operator we may now specify the program for the last Tweety example as

$$\begin{array}{lll}\{\, \mathit{fly}(X) \leftarrow \mathit{bird}(X) \wedge \neg \mathit{ab}(X), & & \\ \;\; \mathit{ab}(X) \leftarrow \mathit{ctxt}\, \mathit{kiwi}(X), & \mathit{ab}(X) \leftarrow \mathit{ctxt}\, \mathit{penguin}(X), & \\ \;\; \mathit{bird}(\mathit{tweety}) \leftarrow \top, & \mathit{bird}(\mathit{jerry}) \leftarrow \top & \}.\end{array}$$

The corresponding ground program is

$$\begin{array}{lll}\{\, \mathit{fly}(\mathit{tweety}) \leftarrow \mathit{bird}(\mathit{tweety}) \wedge \neg \mathit{ab}(\mathit{tweety}), & & \\ \;\; \mathit{fly}(\mathit{jerry}) \leftarrow \mathit{bird}(\mathit{jerry}) \wedge \neg \mathit{ab}(\mathit{jerry}), & & \\ \;\; \mathit{ab}(\mathit{tweety}) \leftarrow \mathit{ctxt}\, \mathit{kiwi}(\mathit{tweety}), & \mathit{ab}(\mathit{tweety}) \leftarrow \mathit{ctxt}\, \mathit{penguin}(\mathit{tweety}), & \\ \;\; \mathit{ab}(\mathit{jerry}) \leftarrow \mathit{ctxt}\, \mathit{kiwi}(\mathit{jerry}), & \mathit{ab}(\mathit{jerry}) \leftarrow \mathit{ctxt}\, \mathit{penguin}(\mathit{jerry}), & \\ \;\; \mathit{bird}(\mathit{tweety}) \leftarrow \top, & \mathit{bird}(\mathit{jerry}) \leftarrow \top & \}.\end{array}$$

Weakly completing the ground program we obtain

$$\begin{array}{lll}\{\, \mathit{fly}(\mathit{tweety}) \leftrightarrow \mathit{bird}(\mathit{tweety}) \wedge \neg \mathit{ab}(\mathit{tweety}), & & \\ \;\; \mathit{fly}(\mathit{jerry}) \leftrightarrow \mathit{bird}(\mathit{jerry}) \wedge \neg \mathit{ab}(\mathit{jerry}), & & \\ \;\; \mathit{ab}(\mathit{tweety}) \leftrightarrow \mathit{ctxt}\, \mathit{kiwi}(\mathit{tweety}) \vee \mathit{ctxt}\, \mathit{penguin}(\mathit{tweety}), & & \\ \;\; \mathit{ab}(\mathit{jerry}) \leftrightarrow \mathit{ctxt}\, \mathit{kiwi}(\mathit{jerry}), \vee \mathit{ctxt}\, \mathit{penguin}(\mathit{jerry}), & & \\ \;\; \mathit{bird}(tweety) \leftrightarrow \top, \quad \mathit{bird}(\mathit{jerry}) \leftarrow \top & & \}.\end{array}$$

whose least model maps *kiwi*(*jerry*), *kiwi*(*tweety*), *penguin*(*jerry*) and *penguin*(*tweety*) to unknown, *ab*(*jerry*) and *ab*(*tweety*) to false (thanks to the *ctxt* operator), and *bird*(*jerry*) and *bird*(*tweety*) as well as *fly*(*jerry*) and *fly*(*tweety*) to true. Reasoning with respect to this least model allows to conclude that *Jerry flies* by *default* without even considering the exceptional cases of birds which do not fly.

If we learn, for example, that

Jerry is a penguin,

then the situation changes again. Our background knowledge is extended to

$$\begin{array}{lll}\{\, \mathit{fly}(X) \leftarrow \mathit{bird}(X) \wedge \neg \mathit{ab}(X), & & \\ \;\; \mathit{ab}(X) \leftarrow \mathit{ctxt}\, \mathit{kiwi}(X), & \mathit{ab}(X) \leftarrow \mathit{ctxt}\, \mathit{penguin}(X), & \\ \;\; \mathit{bird}(\mathit{tweety}) \leftarrow \top, & \mathit{bird}(\mathit{jerry}) \leftarrow \top & \\ \;\; \mathit{penguin}(\mathit{jerry}) \leftarrow \top & & \}.\end{array}$$

The least model of the weak completion of this program maps *kiwi*(*jerry*), *kiwi*(*tweety*) and *penguin*(*tweety*) to unknown and *penguin*(*jerry*), *bird*(*jerry*) and *bird*(*tweety*) to true. Consequently, *ab*(*jerry*) is mapped to true, whereas *ab*(*tweety*) is still mapped to false. Hence, *fly*(*tweety*) is mapped to true, whereas *fly*(*jerry*) is mapped to false. Thus, we can conclude that *Tweety flies, but Jerry does not fly*.

Let us consider the following characterizations about kiwis and penguins:

If a bird has feathers like hair then it is likely to be a kiwi.
If a bird is black and white then it is likely to be a penguin.

Consequently, let us extend the above program with the following clauses:

$$\{\ kiwi(X) \leftarrow featherslikeHair(X) \wedge \neg ab_k(X), \qquad ab_k(X) \leftarrow \bot,$$
$$penguin(X) \leftarrow blackAndWhite(X) \wedge \neg ab_p(X), \qquad ab_p(X) \leftarrow \bot\ \}.$$

Let us assume that we observe

Tweety has feathers like hair and cannot fly.

The most plausible explanation in the context of knowing that *Tweety has feathers like hair*, is, that *Tweety is a kiwi*, and not, that *Tweety is a penguin*.

Hypothesis 5 Humans prefer certain explanations depending on the context.

In order to provide a preference among explanations, [10] provides an extension for abduction under the WCS, which refines the notion of strong dependency among literals in a program. Strong dependency among literals is given when the dependency path is not via a literal within the *ctxt* operator. Accordingly, only abducibles that strongly depend on the observation, are those that are allowed to serve as explanations. For the above observation *Tweety has feathers like hair*, can be explained by itself, which maps $kiwi(tweety)$ to true and at the same time serves as explanation for *Tweety cannot fly*.

5 Obligation Versus Factual Conditionals

This example is taken from [21] with minor modifications. Consider the background knowledge:

If it rains then the roofs are wet and she takes her umbrella.

In fact, these are two conditionals with the same antecedent which can be represented by the program

$$\{wet\_roofs \leftarrow rain \wedge \neg ab_w,\ ab_w \leftarrow \bot,\ umbrella \leftarrow rain \wedge \neg ab_u,\ ab_u \leftarrow \bot\},$$

where *wet_roofs* denotes that *the roofs are wet*, *rain* denotes that *it rains*, *umbrella* denotes that *she takes her umbrella* and ab_w and ab_u are abnormality predicates. Weakly completing this program we obtain

$$\{wet\_roofs \leftrightarrow rain \wedge \neg ab_w,\ ab_w \leftrightarrow \bot,\ umbrella \leftrightarrow rain \wedge \neg ab_u,\ ab_u \leftrightarrow \bot\},$$

whose least model maps ab_w and ab_u to false and all other atoms to unknown.

Given this background knowledge we would like to evaluate the four conditionals

If the roofs are not wet then it did not rain.
If she did not take her umbrella then it did not rain.
If the roofs are wet then it did rain.
If she took her umbrella then it did rain.

To this end we apply minimal revision followed by abduction as presented in Sect. 3. All conditionals are mapped to true. Because the antecedents of the conditionals are unknown given the least model of the weak completion of the background knowledge, abduction is applied yielding the minimal explanation $\{rain \leftarrow \bot\}$ in case of the first and second conditional and $\{rain \leftarrow \top\}$ in the third and fourth conditional. Hence, the consequences of the four conditionals will be true.

This is a bit surprising. Consider the first two conditionals, where the negative consequence of the background knowledge is abductively affirmed, i.e. the consequence is denied. It is not clear that given the information that *the roofs are not wet*, humans would conclude *it did not rain* in the same way as they would conclude *it did not rain* in case the given information was *she did not take her umbrella*. A similar difference holds regarding the third and fourth conditionals, where the consequent of the background knowledge is confirmed. Given the information that *she took her umbrella*, it is again not clear that humans would conclude *it did rain* in the same way as they would conclude *it did rain* given the information that the *roofs are wet*.

5.1 Obligation and Factual Conditionals

It appears that the two conditionals stated as background knowledge should be semantically interpreted in two different ways. Such a semantic interpretation will be developed in the remainder of this section. Consider the conditionals

if it rains then the roofs are wet

and

if it rains then she takes her umbrella.

The consequence of the first conditional is obligatory. We cannot easily imagine a case, where the antecedent is true and the consequence is not. On the other hand, we can easily imagine a situation, where the antecedent of the second conditional is true and the consequence is not. The consequence of the second conditional is not causally obligatory. We will call the first conditional an *obligation conditional* and the second one a *factual conditional*.

As explained in [3], a conditional whose consequence is denied is more likely to be evaluated to true if it is an obligation conditional. This happens because for this type of conditional there is a forbidden or unlikely possibility where antecedent and not consequent happen together and, in this case, where the consequent is known to be false, it cannot be the case that the antecedent is true as, otherwise, the forbidden possibility is violated. Thus, not antecedent is concluded. Because in the case of a factual conditional this forbidden possibility does not exist, a conditional whose consequence is denied should be evaluated as unknown.

Hypothesis 6 Humans may classify conditionals as obligation or factual conditionals.

This is an informal and pragmatic classification. It depends on the background knowledge and experience of a subject as well as on the context in which a conditional is stated.

One should observe that the conditional

Birds usually fly

considered in Sect. 4 is often classified as a factual conditional. For most humans, its consequence is not obligatory. They can imagine cases, where its antecedent *X is a bird* is true and its consequent *X can fly* is not.

5.2 Necessary and Non-necessary Antecedents

The antecedent of the first conditional in the background knowledge is *necessary*. The consequent cannot be true unless the antecedent is true. The antecedent of the second conditional in the background knowledge does not appear to be necessary. There are many different reasons for taking an umbrella like, for example, that the sun is shining. The antecedent of the second conditional is *non-necessary*.

Hypothesis 7 Humans may classify antecedents as necessary or non-necessary.

The classification is informal and pragmatic. It depends on the background knowledge and experience of a subject as well as on the context in which the condition is stated.

5.3 Reasoning with Obligation and Factual Conditionals

Obligation and factual conditionals are represented by programs as before. Thus, the program representing the background knowledge remains unchanged. However, the semantics of conditionals is taken into account by modifying the set of abducibles for a given program. As usual, undefined relations may be abduced. In addition we allow to abduce expressions of the form $A \leftarrow \top$ if A is the head of a conditional with non-necessary antecedent. The consequent A may be true even if the antecedent is not true. And we allow to abduce defeaters of abnormalities if the abnormality occurs in the body of a factual conditional. (One should observe that this is consistent with the use of defeaters in the examples discussed in Sect. 4.) In our example we obtain the set

$$\{rain \leftarrow \top,\ rain \leftarrow \bot,\ ab_u \leftarrow \top,\ umbrella \leftarrow \top\}$$

of abducibles.

Given this modified set of abducibles we may now evaluate the four conditionals using the principle *minimal revision followed by abduction*. Evaluating the conditional

if the roofs are not wet then it did not rain

yields the same result as before because the explanation $\{rain \leftarrow \bot\}$ is the only minimal explanation for the observation that the *roofs are not wet*. Hence, the conditional is true.

Evaluating the conditional

if she did not take her umbrella then it did not rain

we now find two minimal explanations for the observation that *she did not take her umbrella*, viz. $\{rain \leftarrow \bot\}$ and $\{ab_u \leftarrow \top\}$. If we add the first explanation to the background knowledge, then *rain* is mapped to false. However, if we add the second explanation to the background knowledge, then *rain* remains unknown. Reasoning skeptically the conditional is evaluated as unknown.

Evaluating the conditional

if the roofs are wet then it did rain

yields the same result as before because the explanation $\{rain \leftarrow \top\}$ is the only minimal explanation for the observation that the *roofs are wet*. Hence, the conditional is true.

Evaluating the conditional

if she did take her umbrella then it did rain

we now find two minimal explanations for the observation that *she did take her umbrella*, viz. $\{rain \leftarrow \top\}$ and $\{umbrella \leftarrow \top\}$. If we add the first explanation to the background knowledge, then *rain* is mapped to true. However, if we add the second explanation to the background knowledge, then *rain* remains unknown. Reasoning skeptically the conditional is evaluated as unknown.

According to this classification, [21] can solve the abstract as well as the social case of the selection task: In the abstract case the conditional is interpreted as a factual conditional with necessary antecedent and in the social case the conditional is interpreted as an obligation conditional with non-necessary antecedent.

6 Conclusions

We have shown that different human reasoning episodes can be adequately modeled with abductive reasoning. Interestingly, different forms of abduction seem to appear in the different cases. As exemplified by the suppression task, reasoning within conditionals requires skeptical abduction. On the other hand, when modeling reasoning within conditionals and counterfactuals, a form of minimal revision prior to abduction seems to occur. Finally, the last two cases show that the plausibility of explanations can be determined by the context and by the types of conditions and conditionals.

This leads us to two related observations: (1) From a complexity point of view, computing skeptical conclusions is quite expensive [11, 14], and in case the set of abducibles become larger, skeptical reasoning appears to be infeasible. (2) From a cognitive perspective, it does not seem adequate that humans consider all explanations as equally plausible, but rather, that they only consider a *relevant* subset *bounded* by the context, other observations and the type of conditional.

In order to investigate how these explanations have to be bounded, we will need to know more about how humans reason. In general, we need findings from other psychological experiments that can tell us whether the underlying assumptions of the *Weak Completion Semantics* are adequate. In particular, whether humans do *always* prefer skeptical over credulous abduction and how exactly do they apply *revision* when reasoning about counterfactuals.

From our perspective the following questions are desirably worth evaluating.

- *Do humans reason with multi-valued logics and, if they do, which multi-valued logic are they using? Can an answer 'I don't know' be qualified as a truth value assignment or is it a meta-remark?*
- *What do we have to tell humans such that they fully understand the background information?*
- *Do humans apply abduction and/or revision if the condition of a conditional is unknown and, if they apply both, do they prefer one over the other? Do they prefer skeptical over credulous abduction? Do they prefer minimal revisions?*
- *How important is the order in which multiple conditions of a conditional are considered?*
- *Do humans consider abduction and/or revision steps that turn an indicative conditional into a subjunctive one?*

We surmise that humans do reason with a third truth value; we have shown that the *Suppression* and the *Selection Tasks* can be adequately modeled under *Weak Completion Semantics* and, moreover, in these tasks skeptical abduction had to be applied [8]. However, we also believe that they only consider certain explanations as plausible. These are determined by the context, the additional observations and the type of conditional in consideration. Finally, we venture that minimal revision followed by abduction are applied if the conditions of a conditional are *unknown*.

Acknowledgements Luís Moniz Pereira acknowledges support from FCT/MEC NOVA LINCS PEst UID/CEC/04516/2019.

References

1. Adams, E.W.: Subjunctive and indicative conditionals. Found. Lang. **6**(1), 89–94 (1970)
2. Byrne, R.M.J.: Suppressing valid inferences with conditionals. Cognition **31**, 61–83 (1989)
3. Byrne, R.M.J.: The Rational Imagination: How People Create Alternatives to Reality. MIT Press, Cambridge, MA, USA (2007)
4. Clark, K.L.: Negation as failure. In: Gallaire, H., Minker, J. (eds.) Logic and Data Bases, vol. 1, pp. 293–322. Plenum Press, New York, NY (1978)
5. Costa, A., Dietz Saldanha, E.-A., Hölldobler, S., Ragni, M.: A computational logic approach to human syllogistic reasoning. In: Gunzelmann, G., Howes, A., Tenbrink, T., Davelaar, E.J. (eds.) Proceedings of the 39th Annual Conference of the Cognitive Science Society (CogSci 2017), pp. 883–888. Cognitive Science Society, Austin, TX (2017)
6. Dietz, E.-A., Hölldobler, S., Höps, R.: A computational logic approach to human spatial reasoning. In: IEEE Symposium Series on Computational Intelligence (SSCI 2015), pp. 1627–1634. IEEE (2015)
7. Dietz, E.-A., Hölldobler, S., Pereira, L.M.: On conditionals. In: Gottlob, G., Sutcliffe, G., Voronkov, A. (eds.) Global Conference on Artificial Intelligence (GCAI). Epic Series in Computing. EasyChair (2015)
8. Dietz, E.-A., Hölldobler, S., Ragni, M.: A computational logic approach to the suppression task. In: Miyake, N., Peebles, D., Cooper, R.P. (eds.) Proceedings of the 34th Annual Conference of the Cognitive Science Society, pp. 1500–1505. Cognitive Science Society (2012)
9. Dietz, E.-A., Hölldobler, S., Ragni, M.: A computational logic approach to the abstract and the social case of the selection task. In: Proceedings of the 11th International Symposium on Logical Formalizations of Commonsense Reasoning (COMMONSENSE) (2013)
10. Dietz Saldanha, E.-A., Hölldobler, S., Pereira, L.M.: Contextual reasoning: usually birds can abductively fly. In: Proceedings of 14th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR), pp. 64–77. Springer (2017)
11. Dietz Saldanha, E.-A., Hölldobler, S., Philipp, T.: Contextual abduction and its complexity issues. In: Proceedings of the 4th International Workshop on Defeasible and Ampliative Reasoning (DARe) co-located with the 14th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR). CEUR Workshop Proceedings, vol. 1872, pp. 58–70. CEUR-WS.org (2017)
12. Hölldobler, S.: Logik und Logikprogrammierung 1: Grundlagen. Kolleg Synchron, Synchron (2009)
13. Hölldobler, S., Kencana Ramli, C.D.P.: Logics and networks for human reasoning. In: Alippi, C., Polycarpou, M.M., Panayiotou, C.G., Ellinas, G. (eds.) International Conference on Artificial Neural Networks (ICANN), Part II. Lecture Notes in Computer Science, vol. 5769, pp. 85—94. Springer, Heidelberg (2009)
14. Hölldobler, S., Philipp, T., Wernhard, C.: An abductive model for human reasoning. Logical Formalizations of Commonsense Reasoning (Commonsense 2011). AAAI Spring Symposium Series Technical Reports, pp. 135–138. Cambridge, MA (2011)
15. Kakas, A., Kowalski, R., Toni, F.: Abductive logic programming. J. Logic Comput. **2**(6), 719–770 (1992)
16. Kleene, S.C.: Introduction to Metamathematics. North-Holland, Amsterdam (1952)
17. Lloyd, J.W.: Foundations of Logic Programming. Springer, New York, NY, USA (1984)
18. Łukasiewicz, J.: O logice trójwartościowej. Ruch Filozoficzny **5**, 169–171 (1920). English translation: On three-valued logic. In: Łukasiewicz, J., Borkowski, L. (eds.) Selected Works, pp. 87–88. North Holland, Amsterdam (1990)
19. Pearl, Judea: Causality: Models, Reasoning, and Inference. Cambridge University Press, New York, USA (2000)
20. Reiter, Raymond: A logic for default reasoning. Artif. Intell. **13**(1–2), 81–132 (1980)

21. Dietz Saldanha, E.-A., Hölldobler, S., Rocha, I.L.: Obligation versus factual conditionals under the weak completion semantics. In: Hölldobler, S., Malikov, A., Wernhard, C. (eds.) YSIP2—Proceedings of the Second Young Scientist's International Workshop on Trends in Information Processing, Dombai, Russian Federation, 16–20 May 2017. CEUR Workshop Proceedings, vol. 1837, pp. 55–64. CEUR-WS.org (2017)
22. Stenning, K., van Lambalgen, M.: Semantic interpretation as computation in nonmonotonic logic: the real meaning of the suppression task. Cogn Sci **6**(29), 916–960 (2005)
23. Stenning, K., van Lambalgen, M.: Human Reasoning and Cognitive Science. A Bradford Book (2008)