Table of Contents

Essays

In memoriam
Paolo Parrini

ISSN (print) 2281-8618-ETS

Focus
Ian Hacking and the Historical Reason of the Sciences

Past Present
Le rationalisme appliqué

Essays

Virtual trust: Persuasion in social media

Marta Vero

Abstract: With this paper I aim to address the topic of trust on the Internet by associating it with the invention of classical rhetoric in Aristotle's thought. In particular, I will ask whether the influencer can be conceived as the addressee of virtual trust (trustee), by recurring to the *Rhetoric.* Aristotle had already discovered the close nexus between trust and bodily-oral performance. This connection was indeed one of the fundamental tasks of the classical *rhetor.* I claim that this Aristotelian nexus has been maintained through modernity and employed in the Web 2.0. In conclusion, I refer to Instagram celebrities ("influencers") to examine their use of bodily performance to promote purchases or ideas, and to gain the trust of users in order to gain real leverage over their on- and offline lives.

Keywords: trust; Internet; Aristotle, *Rhetoric*; persuasion; influencers

Trust has always been a fundamental keyword of our social life. Although human societies have gone through innumerable transformations, trust has never lost its primary role in every kind of social formation. To appear trustworthy, or to perform virtues in order to be trusted by others is something essential to our social life (Faulkner *et al.* 2017: 3). It follows that trust must be important even for the "artificial" occasions which constitute our social life. This is the case with the Internet. But how can the Internet – an immaterial, virtual reality – be a location of interpersonal relationships based on trust? How can the Internet be intended as a safe environment, where trust can flourish?

These questions are the starting point of this essay,[1] in which I argue that

[1] This paper is the result of a study which I conducted in 2019, thanks to my participation in the research project of the University of Pisa (PRA) "Ethics, Science, Democracy", coordinated by Prof. Adriano Fabris. The main focus of the project was the definition of trust, seen in its "prismatic shape." Essential to my research and to my decision to concentrate on the trust-persuasion nexus in social media were two issues of the journal "Teoria" (both entitled "The Prismatic Shape of Trust"; see "Teoria" 39/2019/1 and "Teoria" 40/2020/1) and the (homonym) cycle of seminars (held in Pisa in the same year), which were conceived in the context of this PRA. For these reasons, I would like to thank the entire scientific committee of the project, and above all Prof. Adriano Fabris, for the philosophical

philinq IX, 1-2021, pp. 9-28
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.270
Submitted: December 2019
Accepted: September 2020

the presence of influencers on the Internet can be understood through a reference to trust and ancient oratory.

1. *Can we trust people on the Internet?*

As Niklas Luhmann (1973; Eng. tr. 1979) said, trust is a basic human skill. It allows us to live, putting faith in the matter-of-fact of nature. It also permits us to determine "the correct and appropriate starting point for the derivation of rules" (4) so that we can orient ourselves in a world otherwise shaped by chaos alone. In short, trust is essential to human action. According to Luhmann, we could claim that trust is a necessary skill for both the natural and the artificial existence of human beings. We have to trust ourselves in order to carry on in a natural, obscure world, which completely hides from us every future event and consequence. In the same breath, we are supposed to trust others and to gain others' trust so as to establish good relationships with them despite the high complexity of the reality of social human life. Complexity, a very abstract attribute of every kind of environment, has to do with our social dimension inasmuch as the increase in human interactions displays the potential infinitude of human possibilities. In other words: a good, functioning social system relies on trust-based relationships. The mechanism of a very complex structure, grounded in connectivity and human exchange, depends on the possibility of people to gain trust inside of it. A complex reality needs to be simplified in trust-based relationships (between trustees and trustors).

It follows that the need for trust must grow proportionally to the openness of our world. In a globalized world, trust becomes, if possible, even more essential for human acting. A world which displays a globally extended net of interactions must deal with a tremendous increase in complexity and with the following escalation of uncertainty. Consequently, a social system which is capable of growing more complex and opening itself to a global perspective must deal with its exposure to uncertainty in order to survive. This means that trust is something needed in such a globalized system, that it must flow inside of it, in order to lubricate its intricate junctures.

All this becomes even more evident when we consider the most extended, interactive social infrastructure of the contemporary world: I am referring, of course, to the Internet. In fact, the Internet has built a serviceable network of communications and information, nowadays actively used by more than half of the world's population, which has increased its reach with enormous speed.

inspirations that this project gave to me. I would also like to thank Prof. Leonardo Amoroso for his support, and Irene Battaglia for her valuable suggestions.

A quick overview of the scientific literature on this topic from the past two decades reveals that trust has always been essential to the Internet as a high-complexity system; nonetheless, it has also been a significant problem. The general concern about trust in virtual communication was one of the leitmotifs of a book (Castelfranchi *et al.* 2001) conceived in 1998 on the occasion of the workshop *Deception, Fraud and Trust in Virtual Agents*. The introduction to this volume makes clear the reasons for this concern: even though the intention was already for the Internet to become a sort of golden goose for the market, it was also clear that its potential could only be exploited if trust flowed in it. The Internet could work as an interaction-based network only if it could sustain trust-based relationships. Or rather, in order for people who wanted to use the Internet as a means for profit, it was necessary to excogitate ways to gain trust from other users. At the end of the 1990s, when the current growth and distribution of the Internet was something still yet to come, the link between trust and the new electronic communication system, in which people exchange their data and share commercial transactions, was already undeniable.

The problem with the possibility of imagining trust-based relationship inside the Internet was broad. Its newness and unknown possibilities were only a part of the question. There was, indeed, another aspect that made the Internet something untrustworthy: its un-corporeality. It is difficult, says Castelfranchi, "to establish trust in an electronic network environment where you usually never meet your partner face to face" (XVIII). Anonymity and bodiless interaction appeared to be unavoidable attributes of this virtual reality. The Internet seemed an artificial, alternative and immaterial place where humans could interact merely with the help of written communication. This is why, in Castelfranchi's perspective, it may be useful to improve the "external factors of trust" (XIX) of the virtual world (that is to say, to make the Internet a reliable environment) to remedy its anonymity or the difficulty of gaining "internal factors of trust" on the Internet – that is, to verify the ability, knowledge, motivation, morality, etc. of those who attempt to become the addressees of trust on the Internet (trustees). The author's point lies in the complementarity of the external and internal factors of trust: "when I strongly trust X [internal trust], then I can accept a less safe and reliable environment [external trust]" (XIX), and *vice versa*. According to the author, one must consider the complementarity between the *external* factors that make reliable the whole environment, and the *internal* factors that make people trust their interactor or mediating agent – a physical person. We have to consider that, in a virtual network based on anonymity and bodiless interaction, as Web 1.0 was at the time, both internal and external factors of trust are very difficult to obtain. As we will see in this paper, Web 2.0 has allowed virtual trustees to overcome the difficulty of gain-

ing "internal" factors of trust and has accordingly allowed the Internet to appear to be a reliable infrastructure.

The postulate that the Internet's anonymity provoked the absence of virtual trust led to the perspective which has been dominant for the past two decades, namely that the Internet was building a different kind of reality, one which was opposed to the material, phenomenal one; a bodiless reality, where trust could not flourish anyway. The strong antithesis between virtual and material reality was, according to Philip Pettit (2004), fundamental. Pettit claimed that authentic trust could not be provoked through the Internet because a virtual reality could not satisfy the conditions for trust. Trust must be based on a dynamic interaction between two agents. It must be addressed to real people (*persons*) and it must be provoked through a *performance*, in a certain, recognizable way. Like any social infrastructure, the Internet itself cannot be *trusted*: it is merely something we can *rely* on. Reliance is indeed something very different from trust: it is a static, intentional, unidirectional human activity, which can only be addressed to a soulless object and motivated by use. According to Pettit, the virtual interaction which takes place on the Internet does not involve real persons. If anything, this kind of interaction takes place between *personas*, i.e. virtual presences, whose faces I "cannot read," who are more similar to ghosts than to real persons: "I cannot read the face of such a contact; the person is a spectral, not a bodily presence in my life" (118). Therefore, the untrustworthiness of *personas* on the Internet rests on a strong separation between virtual and material reality, which relies for its part on the absence or presence of the bodily dimension. In the phenomenal world, the attribution of trust is justified by the fact that I can *see* or *feel* the bodily presence of someone. Their bodily nature and physical proximity (Løgstrup 1956; Eng. tr. 1997) allow me to maintain an actual, dynamic relationship with them, which can also involve trust. To the same extent that it is impossible to trust a robot, as it is any artificial agent or AA (Fossa 2019), an avatar counts as a *persona*; a static picture in an email inbox cannot in any way be considered a trustee, an addressee of authentic trust. According to Pettit, in a virtual world ruled by disembodiment, "we all wear the ring of Gyges" (2004: 118-119). The question is: if an unknown, spectral avatar with all the power and privilege of anonymity cannot be trusted in any way, does it follow that virtual trust is impossible *tout court*?

In 2011, Charles Ess (*et al.* 2011) assumed that the duality between virtual and material reality was the main reason for skepticism towards the possibility of trust-based relationships on the Internet. If we think back to earlier computer-mediated interfaces, such as MUD or MOO, we find that the strong dichotomy between the virtual and the real simply retraced the Cartesian dualism of

mind and body (3). As we saw before, the early Internet seemed to establish a powerful duality between on- and offline dimensions, of a disembodied reality *versus* a physical, material one. The claim of a drastic distinction between the virtual and the real could even encourage an image of the Internet as a "minds only realm" (5), capable of erasing the differences and discriminations (based on ethnicity, gender, religion) which were apparently so hard to overcome in the real, offline world. A wide range of "techno utopian" perspectives have been based on this kind of dualism, from Barlow's cyber-libertarianism to Gibson's idea of cyber-freedom in the novel *Neuromance*. Ess (6-7) also noted that feminist perspectives on cyberspace (Braidotti 1996; Stone 1991; Kember 2003) made it possible to think about the risks of such a cartesian "trick" by anticipating the dependence of the virtual on the real and thinking of the online dimension as an extension of the offline dimension. In general, this pronounced dualism led to the illusion that the Internet could establish a radically different reality, without any dependence on the material world. On the Internet, humans could experience a bodiless identity.[2]

This inevitably means that existence online seems to have lost the essential feature of human embodied life: its materiality. We can already glimpse something which will become clearer in section two, namely the strong connection between trust, bodily performance and orality. The duality between the virtual and the material that dominated early Internet images led to the thought (or perhaps to the hope) that a *disembodied* reality could mean an *alternative* reality, a place of unlimited freedom, where users/*personas* could lose their material vulnerability – and, in this way, the possibility to build trust-based relationships.

The process of overcoming the "Cartesian" dualism which ruled theorists' comprehension of Web 1.0 went hand in hand with the technological development of the Internet as a communication system. Furthermore, the increasing immersivity of online experiences, the standing-out of a "virtual *continuum*" (Milgram *et al.* 1994: 2), permitted also by the rapid modernization of interfaces, has allowed the boundaries between the virtual and the real to be demolished. The so-called Web 2.0 is a deeply interactive reality which finds its essential predicate in offering the possibility to share content with other people, as well as to modify and reproduce it. In Carusi's view (2011), this opportunity to interact by manipulating content in every moment contributed to a very different perception of the virtual world. In the 2.0 era, a radical dualism between the virtual and the real no longer makes sense because our online experience has become inseparable from our offline life. Carusi speaks of an

[2] According to Kember (2003), the defeat of this kind of dualism between body and mind, operated by Web 2.0, is one of the reasons for the stagnation of cyberfeminism.

"interworld" (115) between physical and virtual reality as the ideal place where we conduct our social life.

Myskja's (2011) argument takes a similar direction, in which the virtual dimension of human existence is seen as an artificial one. However, the fact that something is artificial does not mean that is not at all real. This clever insight is expressed by the analogy of virtual reality with Kant's use of fictionality, or the "philosophy of 'as if'" (121). This means that, on the Internet, as in every artificial (artistic, ritual and so on) context, we are all aware of the fictional character of the virtual reality we are helping to build. This is why in Myskja's perspective, "virtual reality is a form of art in the sense described by Kant. We know that this is fictional, but we must treat it as a reality in order to make it work the way we intend. It is a deception where nobody is deceived" (131). Claiming the fictionality of the Internet does not make it less real than material life, just as is the case with regard to fictional places such as theatres, or stadiums of ancient oratory. Just as in art or other situations which require specific performances, fictionality leads humans to acquire specific behaviors, manners or attitudes to appear more attractive to others, in order to be trusted. With the arrival of Web 2.0, the social centered virtual reality known as the Internet ceased to be an untrustworthy, dangerous space; it had begun to become increasingly real. In other words, the need to imagine the Internet as a dimension which is completely alternative to the material one was decreasing. The Internet became a fictional part of reality where we could assume several different behaviors in accordance with the strategies adopted to gain the trust of others.

When Myskja's paper was published, the Internet was still an almost disembodied environment. This was, like before, the major factor contributing to the untrustworthiness of virtual interactions between humans. However, Web 2.0 was on the rise and was constantly implementing a means to overcome the boundaries between the virtual and the real world and, in this way, to allow itself to sustain trust relationship – or, as Castelfranchi claimed several years before, to improve its external factors of trust in order to overcome its anonymity. For instance, the diffusion of chat rooms created realities which were still bodiless, but where it became possible to *perform* veracity, as well as many other virtues (Vallor 2010),[3] in order to gain trust online. Former chatrooms showed us that the social rules of human interactions were not completely neglected on the Internet: they needed to be performed *as if* they were in the material world. Just think, as Myskja suggests, of interactive online games, such as *Second Life* (2011: 132) or *The Sims,* where users can interact by

[3] According to Hawhee (2004), the nexus between the capacity to perform virtues, agonism and persuasion was at the core of Ancient Greek rhetoric.

displaying behaviors which are very similar to behaviors performed in the real world. The Internet became a place in which people could meet, discuss, fall in love, become friends, trust each other; all of this, without seeing each other. It follows that the supposed inability of the Internet to sustain trust-based relationships must be brought into question. We saw before that the reason theorists assumed that one could not gain a "virtual" kind of trust was the uncorporality of the Internet as an infrastructure grounded in connectivity. This was also why, paradoxically, the Internet needed to overcome its anonymity in order to allow trust to flow inside it and to simplify its complexity.

To summarize, we could say that the enormous potential of the Internet can only be realized with a constant reference to trust. Since the Internet is a very complex and unknown social system, in order for it to become a reliable infrastructure, it must create factors which allow users to trust each other mutually. Unfortunately, gaining trust without being seen, and only with the help of written messages, is very difficult. When reviewing the scientific literature on this topic from the last two decades, it becomes clear that the arrival of Web 2.0 provided several tools to induce trust despite the anonymity of the Internet. Namely, the invention of social networking permitted the introduction of body-like behaviors into virtual communication, which made computer-mediated interactions more similar to the oral ones – that is, more real. One could also argue that the evolution of the Internet consisted in a progressive promotion of the bodily dimension: passing over the strong dichotomy between the phenomenal and the virtual by promoting oral communication and bodily performance enables the former avatars to become actual trustees. The more the Internet can reproduce a physical proximity with its virtual means, the better it is able to promote trust relationships between users; the more the users are able to place trust in their mutual virtual interactions, the better the Internet's (economic, social, political) potential can be expressed. Last but not least, the possibility to persuade one or more unknown persons and consequently to become a trustee is greater when the infrastructure we are relying on allows us to show ourselves in our bodily and oral dimension, since performing a certain kind of behavior will generate trust in others.

But why is performing virtues so essential to trust? And how is this linked to a bodily and oral dimension, which seems so essential in a new means of communication?

2. *Trust and persuasion: performing virtues*

To understand the reasons behind the connection between trust and the bodily/oral dimension, especially by reference to the trustees, we can turn

to Aristotle's philosophy. In his *Rhetoric*, the first systematic treatise on rhetorical argumentation, the Greek philosopher assigned to the rhetor a very clear purpose: to persuade his audience. Whatever the genre of the oration was, knowing the *rhetorike techne* had to lead the orator "to see available means of persuasion (*peitho*) in each case" (*Rhet.* 1355b; Eng. tr. 2007: 36). The verb *peithomai* constitutes the most important notion of Aristotle's *Rhetoric* and presents, upon close inspection, a strong relation to trust. In fact, to the Greeks, persuasion was connected to the sphere of trust so firmly that words related to the verb *peithomai* – such as the substantive *pistis* and the adjective *pistos* – can also be translated as "loyal" and "loyalty," "credible" and "credibility," "trustworthy" and "trustworthiness" (Piazza 2008: 17). At first, the nexus between rhetoric and trust appears motivated by the fact that a rhetorical argument addresses itself to the developing of opinions (*doxa*) in the audience. Rhetoric is in fact an *art* (*techne*) which shows a strong analogy (*antistrophia*) to dialectic (*Rhet.* 1354a; Eng. tr.: 30). However, it cannot be defined as a dialectical practice, since it cannot produce certain knowledge (*episteme*) like philosophy does. In the realm of *doxa*, we have no assurance of the falsity or trueness of our reasoning. This is why the audience is inclined to "commit [itself] to trust our own opinions or convictions" (Piazza 2008: 18-19). It means that, in the realm of *doxai*, opinions are malleable: someone who masters the *rhetorike techne* can generate trust in the audience and, in this way, she can try to direct some other's view, attitude, vote, or preference. Rhetoric is, to this extent, the art of trust.

According to Aristotle, the art of rhetoric is founded on three different "technical *pisteis*" (*Rhet.* 1356a; Eng. tr.: 38) or technical arguments which must be known and well used by the speaker and be distinguished from *non-technical* ones: these are *ethos*, *pathos* and *logos*. While understanding what *pathos* and *logos* mean turns out to be quite simple, considering also the wide space that Aristotle assigns to these technical *pisteis* in *Rhetoric,* we cannot say the same about *ethos.* This is one of the most mysterious aspects of *Rhetoric.* Aristotle devotes long chapters of his treatise to the methods of construction of *enthymemes* or to the strategies of provoking specific emotions, depending on certain factors such as age, origins etc. (*Rhet.* II, §§12-17; Eng. tr.: 149-155) and in general on the composition of the audience. Indeed, only a few lines are dedicated to the explanation of the *ethos*, although Aristotle admits that "*ethos* represents the stronger *pistis*"; *ethos* realizes persuasion inasmuch as "the speech makes the speaker *trustworthy* (*axiopistos*)" (1356a; Eng. tr.: 38).[4] The audience is inclined to believe in a certain cause, to be seduced by an ex-

[4] Kennedy translates *axiopistos* as "worthy of credence."

hortation or moved by a eulogy to a greater extent if the speaker can present himself as a trustworthy person.

The importance of *ethos* for persuasion derives from the centrality of the speaker's oral and bodily performance. Aristotle explains on several occasions that *ethos*, as the other technical *pisteis*, must shine through speech itself. Only through her rhetorical performance is it possible for the speaker to appear trustworthy: this is why *ethos* belongs to the technical argumentations. According to Jakob Wisse, *ethos* can be referred to as the "character" (1989: 30-33) of the rhetor,[5] and must express the moral and intellectual qualities (virtues) of the speaker. These virtues, like good sense (*phronesis*), goodwill (*eunoia*) and goodness (*arete*), are fundamental for a speaker to be trusted, because in absence of such (moral) qualities one can doubt that the rhetor is telling the truth (32).[6] That speech has to manifest goodwill, good sense and goodness means that these virtues count for the purpose of persuasion only if they appear through the speech itself, as we said. It follows that the previous reputation of the speaker must be irrelevant for the success of persuasion.

The rhetorical potential of the *pistis* grounded on *ethos* is, in this respect, very close to that of *pathos*. The speaker must establish emotional proximity to the audience, in which the listener "suffers along with the pathetic speaker, even if what he says amounts to nothing" (*Rhet.* 1408a; Eng. tr. 210). The listeners allow the speaker to provoke emotions in them and in this way to influence their process of opinion-making on the basis of the trust that the speaker's *ethos* can create. In this respect, certain authors have spoken of "inactive friendship" (Piazza 2008: 97), a concept that we could relate to the Kantian idea of "as if."[7] In other words, the speaker moves in a fictional space, where she discovers the power of making extensive use of human emotions, verbal strategies and also her own virtues in order to persuade the audience. The rhetors act "as if" they are in the "real" world and employ strategies that can be used in ordinary life. All this, however, happens in the fictional space of an assembly, where a speaker can take advantage of emotions and virtues only if these are technical *pisteis*, contained in an oration. This means that even the speaker's character becomes persuasive only in the fictional space of the oration, of the spoken word (Ong 2012: 31-33; Hawhee 2004). The rhetor is hence an actual performer. The logical element of his speech is naturally so important that Aristotle dedicated the majority of his treatise to the *inventio, elocutio* and *dispositio* of persuasive speech, the branches upon which modern oratory also relies.

[5] See also Di Piazza (2012).

[6] See also Perelman (*et al.* 1958; Eng. tr. 1971); Hawhee (2004).

[7] Piazza refers to *NE* (1167a).

There is no speech which can persuade without being performed. The most powerful ability of the rhetor lies, according to Aristoteles, in her technical capacity to make use of irrational components in rationally constituted speech in order to arouse trust in the audience. This Aristotelian idea is conceived not only against Plato's disdain for rhetoric as a non-philosophical discipline, but it is also an attempt to turn the irrational, socially "magic" aspects of rhetoric,[8] highlighted especially by Gorgia, into something that could emerge in a *logos* assembled by specific rules. Just as the treatise on *Poetics* formulates criteria for evoking *eleos* and *phobos* (*Poet.* 1345b; Eng. tr. 2000: 18) in the audience during a tragic performance, the *Rhetoric* studies the process of arousing trust through the speaker's *ethos* in deliberative, judicial or epideictic assemblies.[9]

At the center of this strong analogy between the Aristotelian arts of rhetoric and poetry lies the *performativity* of the actor – the rhetor. In his treatises, the Stagirite depicts two of the occasions, such as theater and assemblies, which were essential for life in the Greek *polis*. They created a strong emotional cohesion between those citizens who could take part in the assembly or in the tragic performance. Actors and rhetors could manage the moods of the audience so well that they were even able to cause catharsis or induce trust when they wanted. In short, the democratic cohesion of the citizens depended on assemblies and tragic performances: these "fictional" places were indispensable for the political and social life of the Greek *polis*. Reading Aristotle, we note that the oral performer was a kind of orchestra leader in managing emotions and evoking trust in the audience. The *logos* on which Aristotle sets up the *Rhetoric* does not mean, for this reason, any (written) discourse, just as a Sophoclean tragedy could not be intended merely as a written test. The Aristotelian connection between *ethos, pathos* and *logos* points to a wise management of the *spoken word* and, consequently, refers to the ancient Greek centrality of the staging performance.[10]

This comparison shows that the centrality of religious, cultural, judicial and political meetings was founded on the interaction between word and *praxis*. The political sense addressed to these occasions "demonstrates mutuality among performative sites that create their credibility by reiterating familiar patterns of language and actions" (Miller 2007: 60). Ancient performers had to insist on recalling the notion of community. Their speech had to focus on sharing emotions, creating an atmosphere of common feelings, but also on

[8] See De Romilly (1975); Dodds (1951); Butler (1999).

[9] See Miller (2007: 58).

[10] See Bonanno (1997); Hawhee (2004).

recollecting a commonality of virtues and value patterns. In the case of the rhetor, this was essential in order to induce trust, and consequently to persuade the audience. When Aristotle writes that virtues like goodwill, goodness and good sense must be displayed in discourse, he is referring to affirmations and gestures which could both contribute to giving the speaker a virtuous (*Rhet.* 1418b; Eng. tr.: 245) and trustworthy image. Improving the style of oration is equivalent to increasing the capacity to express "emotions" (1408a; Eng. tr.: 210) in response to certain acts. These emotional expressions have to be appropriated, namely, respondent to socially accepted patterns of behavior, in order to expose the credible *ethos* of the speaker. In this way, the speaker becomes trustworthy and therefore able to persuade.

The centrality of the rhetor's performance in Aristotelian *Rhetoric*, since it is produced through spoken and practical features, is something undoubtedly missing in written texts. As we have seen, the art of persuasion is, according to Aristotle, inextricably associated with the dimension of orality. The whole *rhetorike techne*, as Aristotle (and also the Roman Cicero) imagined it, could not exist without the central figure of the orator, seen as a performing figure, as someone who is able to vehiculate messages and feelings by means of gestures, tones and modulations of the spoken word. In order to better understand why the art of persuasion must, from an Aristotelian perspective, be linked to the public figure of the orator and to her capacity to display virtues, it can be useful to briefly consider what the Stagirite says with regard to the purposes of one of the three species (*eide*) of oratory. I am referring particularly to the *symboleutikon*, or deliberative rhetoric, which, according to Aristotle, must be distinguished from judicial as well as demonstrative rhetoric (1358b; Eng. tr.: 48). Among the three genres of oratory, the deliberative discloses the strongest nexus with the practice of *influencing* others' behavior and opinions. Deliberative oratory consists of exhortations and dissuasions and refers to a specific "time": the future, "for whether exhorting or dissuading [the orator] advises about future events" (1358b; Eng. tr.: 48). As the chapter 4 of the first book shows, the field of application of the deliberative oratory is wide. It deals with political, economic, legal exhortations or dissuasions (1359b; Eng. tr.: 53); but it also relies on essentially ethical goals. To become the addressee of trust, the deliberative orator must know the main values of human life. This means that, in order to persuade an audience and to influence people's decisions about the subjects most important to public life in the polis, one must know human virtues, namely one must refer to ethical topics that can be universally shared by the listeners.[11] This is why chapter 5 begins

[11] Some of the ethical topics mentioned by Aristotle in chapter 5 can be more useful for epideictic oratory, while just a few of them can be employed in judicial oratory.

with a mention of happiness, the *skopos* of human life, which the orator must always keep in mind, since from happiness are derived all the goods that make a person trustworthy. Happiness is, according to Aristotle, having "success [in life; *eupraxia*] combined with virtue" or "self-sufficiency [*autarkeia*]" (1360b; Eng. tr.: 57), being in possession of the goods (which are listed through chapter 5) and having the ability to defend them. Aristotle proclaims that deliberative oratory is very much concerned with this definition of *eudaimonia* and with all the goods (*agatha*) and virtues (*aretai*) which are parts of it. But what does it mean to "know" what goods and virtues are? Is it sufficient for the orator to have the ability to define them? In my opinion, in the close nexus between persuasion and happiness as a performance of the possession of goods and virtues lies the reason why public speech, seen in its materiality and proclaimed in front of an audience, is able to influence the listeners' actions.

Provided that happiness describes the ultimate *skopos* of human life, the deliberative orator recognizes her particular aim in the "advantageous"; she does not need to refer to happiness, but must address herself to the specific "means" (1362a; Eng tr.: 61) to reach happiness, since it is these means which constitute the object of a deliberative assembly and which must be displayed in the speech of the orator. In other words, the sources of the persuasive argument for deliberative speech are the goods and virtues as a means of achieving *eupraxia,* and happiness. Virtues, as something that "are necessarily a good, […] productive of good things and matters of action" (1362a: Eng. tr.: 62), are the ultimate and most relevant ethical goods to which the orator must refer. As the Stagirite clarifies, also in a famous passage in *Nicomachean Ethics* (*NE* 1105b-1106a, Eng. tr.: 28-29), virtues are *hexeis*: dispositions toward good actions, the *habitus* we adopt when we have to take a decision and act as good persons (1106a, Eng. tr.: 28).[12] Aristotle's employment of the term *hexis* in this context suggests that virtues, namely the ethical goods which allow people to reach *eupraxia* and in this way to appear happy, are something which has to be embodied and performed. As a matter of fact, the term *hexis* designates a strong connection between the bodily dimension of *arete* and its theoretical

[12] See also: "And the virtues (*aretai*) are necessarily a good: for those having them are well-off in regard to them, and virtues are productive of good things and matters of action […]. To speak of these one by one, the following are necessarily good: happiness […], justice, courage, temperance, magnanimity, magnificence, and similar dispositions (*hexeis*), for they are virtues of the soul" (*Rhet*., 1362b, Eng. tr.: 62). I do not mean here to express an opinion on the so-called inconsistency of Aristotle's depiction of virtues in the *Rhetoric* and in the *Nicomachean Ethics*. I am merely using the notion of *hexis* to underline the performative aspect of the virtue, not to declare that virtue cannot be a *dynamis*. Regarding solving of this Aristotelian "inconsistency," see Allard-Nelson 2001. This "bodily" sense of *hexis*, translatable with the Latin *habitus*, was famously understood by Pierre Bourdieu (1980, Eng. tr. 1990); see also Butler (1999: 116).

meaning. It is used also in *Metaphysics*, 1009b to describe a physical disposition capable of modifying the thought (*phronesis*) (*Metaph.* 1009b; Eng. tr.: 61),[13] building a bridge between the bodily and theoretical dimension of virtue and confirming that *hexis*, "the Greek word for bodily conditions or bodily state, is indistinguishable from habits and practices" (Hawhee 2004, 58). Stating that virtue is a *hexis* and placing it at the core of a treatise on persuasion is equivalent to asserting that virtue must be *embodied* and *performed* by the orator who wants to be trusted. To become a trustee, the orator *must bodily display* her possession of such virtues, namely the means of achieving happiness, which are the primary objects of deliberative oratory.

The *pistis* which is designated to transform this necessity to perform and embody virtues in a technical, systematical way in order to produce a persuasive speech is certainly *ethos.* The close nexus between *ethos* and virtue (and consequently with *hexis*) can be recognized again in the *Nicomachean Ethics*, where Aristotle describes the character as something that absorbs or incorporates virtues by means of habits: "So virtues (*aretas*) arise in us neither besides nature nor before nature, but nature gives us the capacity to exhibit (*dexasthai*) them, perfecting them by means of the character (*dia tou ethous*)"[14] (*NE* 1103a; Eng. tr.: 23). To answer the difficult question of natural origin of the virtues, Aristotle establishes a connection between *hexis* and *ethos.* Here the author seems to suggest that *ethos*, as a character made up of habits, is anything but the material and time-extended exhibition of virtues or, which is the same, that *ethos* consists of the visible embodiment of ethical dispositions, it is the domicile of their incorporation in our everyday life. Our character, interpreted as a performance of *habitus*, is something subject to our interventions – we can learn to be or to appear virtuous for many reasons, including appearing trustworthy – which becomes like a second nature with the passing of time. In *Rhetoric* I, 11, a section dedicated to judicial oratory, the Stagirite claims that, through habits, *ethos* becomes "natural; for habit (*ta ethe*) is something like nature (*physei*)" (*Rhet.* 1370a, Eng. tr.: 87).

The character to which Aristotle refers in *Rhetoric* is therefore that dimension of human being in which dispositions become "ingrained" (Hawhee 2004: 95) in a person's performance to such an extent that they seem to be

[13] The context in which the term is used in *Metaphysics* is indeed very different from that of *Rhetoric.* In this passage Aristotle is referring to Empedocle's doctrine with regard to the production of *metis.* I only mean to underline that the term *hexis* is evident involved with the bodily dimension, as proved by this quick reference to another occurrence of the term. See also Hawhee (2004: 57-58).

[14] Translation modified. Crisp translates: "nature gives us the capacity to acquire them, and completion comes through habituation". I follow Hawhee (2004: 95) when I translate *dexasthai* with "exhibit" and *teleioumenois* with "perfection."

natural; namely, that they are a spontaneous element of a bodily, visible way of acting. The *ethos* in rhetoric thus represents the visible and performative[15] element of speech, which is capable of arousing the social magic of oratory which lies in persuasion and social influencing. Understood as a technical *pistis* of an orator's persuasion strategy, *ethos* consists of reproducing gestures, habits and words which re capable of making the orator appear trustworthy to the audience. One becomes a trustee by performing virtues as goodwill and *eupraxia*. To appear virtuous, good and independent means to be able to influence others' gestures, perhaps to make the exhibition of virtues reproducible among the listeners. What Aristotle claims with regard to deliberative rhetoric is suggestive of why bodily performance is so central in persuasion, and why gaining trust means first of all performing "virtues" or socially accepted patterns of behavior.

3. *A trustee in social media: the influencer*

It is obviously impossible for us to experience the enormous importance of assemblies and theaters to the Greek *polis.* This is something very difficult to reproduce in a chirographic culture, such as the one which has succeeded in modernity. This is why it seems so difficult to trust the authors of written texts.[16] The transformation of the reference community, initially composed of listeners gathered together in a limited audience and then identifiable as an uncountable community of unknown readers, paradoxically arouse the need for trust. Indeed, the idea that the process of constructing opinions could be freed from persuasion strategies thanks to the promotion of written texts was soon revealed to be pure utopia. Every means of communication relies on performative means of persuasion, like the *ethos* discovered by Aristotle: even the authors of written texts must develop strategies to evoke trust (Miller 2007: 74-80).

But how are things with the Internet? As we have seen, the Internet is a means of communication which initially was not seen as being capable of sustaining trust relationships. Things changed with the arrival of "social web," which introduced bodily and oral elements into the virtual reality, initially marked by anonymity, and have allowed people to become the addressees of virtual trust. The innovations brought about by Web 2.0 brought allowed

[15] On the performative as a crucial notion for aesthetics and rhetoric, see Fischer-Lichte, (2004, Eng. tr. 2008).

[16] The spread of printed texts could, for one thing, promote an autonomous reaction to words and, so to say, protect readers from the more dangerous aspects of oral persuasion. As an example, we could take the chronological connection between the diffusion of printed texts and the Lutheran claim to understand God "*sola scriptura*" (Ess 2011: 15).

people to show themselves in a virtual environment without being limited by written texts. Nowadays, on the Internet one can finally show oneself in a complete way. One can be seen in one's body and gestures, in one's voice,[17] in one's everyday life. This return to the centrality of body and orality on the Web can be seen as something that makes persuasion on the Internet easier than before, and can be observed especially in Web 2.0's most unique invention: social networks. Social networks such as Twitter (Fabris 2012; 2015), Facebook, Pinterest etc. have made connectivity their fundamental keyword. In these media, users can express their opinions in the form of brief statements and instantly receive feedback from their followers. Something even more interesting can be noted about social media which make use primarily of images and voice and which have invented a very productive way to reproduce physical proximity: the daily reiteration of videos which show the same person or events from the everyday life of that person. I am referring to blogs and vlog platforms,[18] as well as to social media platforms such as Snapchat, TikTok, Tumblr and especially Instagram. The latter, increasingly widespread among teenagers and young adults, owes its popularity to its "images only" formula. At the beginning, Instagram was a social media platform in which the only content one could share was, precisely, images, eventually accompanied by a very short tagline and hashtags, that is, keywords which allow the posted image to be disseminated widely. This image-only social interaction has had important consequences in our ordinary life: just think of the massive diffusion of *selfie* culture, which is often interpreted as a sign of our "narcissistic" (Wendt 2014; Sheldon *et. al.* 2016) era.

More interesting for our analysis is the latest innovation from Instagram: the introduction of "live stories." This development occurred throughout 2016 and has changed with surprising rapidity not only social networking interactions between single users, but also marketing strategies. This kind of social media concept, based purely on the circulation of images and real-time videos, proves that the current immersive experience of the Internet integrates the bodily dimension and oral communication in a virtual space. The diffusion of Instagram Stories has led in a very short time to the evolution of a commercial figure such as the influencer, which actually already existed within customer services (Grenny *et al.* 2013), and transformed it into an actual "Internet celebrity." Thanks to Instagram Stories, live stories and recently to

[17] See Ong (1971: 296), according to whom there is a progressive tendency of electronic technologies to introduce oral elements beside textual, written ones.

[18] Vlogs, or video blogs, usually take advantage of other platforms, such as Youtube, Vimeo, etc.

IGTV,[19] the promotion of "self-branding" can take advantage not only of the worldwide diffusion of the influencer's image, but also of their voice, gestures, opinions, lifestyle and actions in everyday life. This is how the influencer figure started to become not only a marketing strategy, like the involvement of celebrities in television commercials in the early 1990s, but also a powerful addressee for trust in virtual communication and transactions which take place on the Internet. With the daily use of Instagram Stories, social media influencers can also give testimony to the value of a product inserting it into the frame of their lifestyle.

Upon close examination, the influencer's role is seen to be extremely linked to the notion of trust, exactly because, as the orators of Aristotle's time, it deals with persuading others to change their habits, to acquire goods or *hexeis* that the influencer shows herself to be in possession of. In this way, the fiction upon which the Internet rests as a reliable infrastructure becomes more successful: The Web ceases to seem like a disembodied and scary place, and acquires a realistic, almost "domestic" appearance. Experiences and relationships on the Internet are always "virtual," non-material, but they look increasingly similar to the ones which happen "irl" (in real life). Influencers prove that, given the possibility to continually show performative skills thanks to the virtual exposition of gestures and orations, trust begins to flow massively on the Internet. An influencer can perform and exhibit her "ethos" online almost daily. In every moment of her life, she can profitably pretend to have certain qualities (such as goodness, goodwill, good sense and especially sincerity) to a potentially infinite audience. We know from Aristotle that recurring to *ethos* to appear trustworthy means in particular to exhibit good dispositions or at least the possession of something which is advantageous, something that the listener could desire for herself. For this reason, an influencer *performs* qualities and exhibit goods in order to be trusted by her followers. The strong interactivity promoted by this kind of social media allows followers to establish a real time communication with influencers,[20] which, for its part, allows the influencer to know exactly what kind of audience she has, in order to modulate her use of passions or arguments, as Aristotle suggested.

Within this framework, influencers represent those located at the center of the fictional arena of the Internet; that is to say, they are the addressees of our virtual trust. There is (still) no handbook of influencing like Aristotle's treatise on *Rhetoric*, but we can see that contemporary influencers have learned

[19] A video platform, created by Instagram, "intended to compete against Youtube" (Wikipedia).

[20] Instagram's "direct messages (DM)" function enables users to give immediate, private feedback on Ig-stories and live videos.

his lesson. Performing trustworthiness is the best way to generate trust in an audience composed of unknown people. In order to do that, one must take advantage of their bodily means and, specifically, of an oral register of expression which can reproduce physical proximity to the audience. Overexposure of Internet celebrities on social media allows them to present their behaviors as evidence of their trustworthiness. They can provoke admiration among their followers, behaving as they would on a stage, interacting with the components of their "audience," receiving immediate feedback and interacting with them, "as if" all this was happening in "real" life.

To summarize, the Internet's former absence of a bodily and oral dimension was the element which made it difficult for us to trust each other inside of it. The Internet, being a fictional place, just as Greek theaters or assemblies were, must reproduce dynamics, relations, etc. which allow it to resemble the real world. By integrating the persuasive power of body proximity and oral communication in a virtual reality, social networks needed a figure that could inspire admiration, sincerity, goodwill and other virtues among users; someone who could behave as if the Internet were a real place to live in, in which people can feel real sentiments, such as admiration and trust. This is what influencers are for. They possess the power to condition lifestyles, purchases, and also the opinions and behaviors of social media users. Their work consists in finding a way to be trusted by a globalized audience. Just like the rhetors of Aristotle's time, they try to exhibit goods and perform virtues in order to gain trust.

Marta Vero
marta.vero90@gmail.com
University of Pisa

Bibliography

Allard-Nelson, Susan K., 2001, "Virtue in Aristotle's Rhetoric: A Metaphysical and Ethical Capacity", in *Philosophy & Rhetoric*, 34, 3: 245-259.

Aristotle, *Rhet.*, Eng. tr. by George A. Kennedy, *On Rhetoric: A Theory of Civic Discourse*, 2007², Oxford University Press, Oxford-New York.

Aristotle, *Poet.*, Eng. tr. by Samuel H. Butcher, *The Poetics of Aristotle*, 2000, Pennsylvania State University Press, Pennsylvania.

Aristotle, *NE*, Eng. tr. by Roger Crisp, *Nicomachean Ethics*, 2004, Cambridge University Press, Cambridge.

Aristotle, *Metaph.*, Eng. tr. by C. D. C. Reeve, *Metaphysics*, 2016, Hackett, Indianapolis.

Baier, Annette, 1986, "Trust and Antitrust", in *Ethics*, 96, 2: 231-260.

Bonanno, Maria Grazia, 1997, "All the (Greek) World's a Stage: Notes on (Not Just)

Dramatic Greek Staging", in Edward Lowell Edmunds, Robert W. Wallace, eds., *Poet, Public, and Performance in Ancient Greek: From Homer to the Fifth Century*, John Hopkins University Press, Baltimore: 112-123.

Bourdieu, Pierre, 1980, *Le Sense Pratique*, Les èditions de Minuit, Paris; Eng. tr. Richard Nice, 1990, *The Logic of Practice*, Stanford University Press, Stanford.

Braidotti, Rosi, 1996, "Cyberfeminism with a Difference", in *New Formations*, 29: 9-25.

Butler, Judith, 1999, "Performativity's Social Magic", in Richard Shusterman, ed., *Bourdieu. A Critical Reader*, Blackwell, Oxford: 113-128.

Castelfranchi, Cristiano, Yao Hua Tan, 2001, eds., *Trust and Deception in Virtual Societies*, Springer, Dordrecht.

De Romilly, Jacqueline, 1975, *Magic and Rhetoric in Ancient Greece*, Harvard University Press, Cambridge and London.

Di Piazza, Salvatore, 2012, "Fiducia e argomentazione. Una prospettiva aristotelica", in *Rivista Italiana di Filosofia del Linguaggio*, 6, 3: 41-52.

Dodds, Eric Robertson, 1951, *The Greeks and the Irrational*, University of California Press, Los Angeles.

Ess, Charles, 2011, "Self, Community, and the Ethics in Digital Mediatized Worlds"; Carusi, Annamaria, 2011, "Trust in the Virtual/Physical Interworld"; Myskja, Bjørn, 2011, "Trust, Lies, and Virtuality" in Charles Ess and May Thorseth, eds., *Trust and Virtual Worlds. Contemporary Perspectives*, Peter Lang, New York: 3-31; 103-120; 120-139.

Fabris, Adriano, 2015, *Twitter e la filosofia*, ETS, Pisa.

Fabris, Adriano, 2012, *Etica delle nuove tecnologie*, La Scuola, Brescia.

Faulkner, Paul, Thomas Simpson, ed., 2017, *The Philosophy of Trust*, Oxford University Press, Oxford.

Fischer-Lichte, Erika, 2004, *Ästhetik der Performativen*, Suhrkamp, Frankfurt; Eng. tr. by Saskia Iris Jain 2008, *The Transformative Power of Performance. A New Aesthetics*, Routledge, Oxon.

Fossa, Fabio, 2019, "'I don't trust you, you faker!' On Trust, Reliance and Artificial Agency", in *Teoria*, 39, 1: 63-80.

Grenny, Joseph, Kerry Patterson, David Maxfield, Ron McMillan, Al Switzler, 2013, *Influencer. The New Science of Leading Change*, 2nd ed., McGraw Hill, New York.

Hawhee, Debra, 2004, *Bodily Arts: Rhetoric and Athletics in Ancient Greece*, University of Texas Press, Austin.

Kember, Sarah, 2003, *Cyberfeminism and the Artificial Life*, Routledge, London.

Løgstrup, Knud Ejler, 1956, *Den Etiske Fordring*, Gyldendal, Copenhagen; Eng. tr. by Hans Fink 1997, *The Ethical Demand*, University of Notre Dame Press, Notre Dame.

Luhmann, Niklas, 1973, *Vertrauen,* Enke, Stuttgart; Eng. tr. by Howard Davis, John Raffan, Kathryn Rooney 1979, T. Bauer, G. Poggi, ed., *Trust and Power: Two works by Niklas Luhmann*, Wiley, Chichester.

Milgram, Paul, Funo Kishino, 1994, "A Taxonomy of Mixed Reality Visual Displays", in *IECE Transactions on Information and Systems*, 12, 12: 1321-1329.

Miller, Susan, 2007, *Trust in Texts: A Different History of Rhetoric*, Southern Illinois University Press, Carbondale.

Ong, Walter J., 1971, *Rhetoric, Romance, and Technology: Studies in the Interaction of Expression and Culture*, Cornell University Press, Ithaca-London.

Perelman, Chaim, Lucie Olbrechts-Tytheca, 1958, *La nouvelle Rhetorique. Traité de l'Argumentation*, Presses Universitaires de France, Paris; Eng. tr. by John Wilkinson, Purcell Weaver 1971, *The New Rhetoric: A Treatise on Argumentation*, University of Notre Dame Press, Notre Dame.

Pettit, Philippe, 2004, "Trust, Reliance and the Internet", in *Analyse & Kritik*, 26: 108-121.

Piazza, Francesca, 2008, *La retorica di Aristotele. Introduzione alla lettura*, Carocci, Roma.

Sheldon, Pavica, Katherine Bryant, 2016, "Instagram: Motives for its use and relationship to narcissism and contextual age", in *Computers in Human Behavior*, 58: 89-97.

Stone, Allucquère Rosanne "Sandy", 1991, *Cyberspace: First Steps*, MIT Press, Cambridge MA.

Vallor, Shannon, 2010, "Social networking technology and the virtues", in *Ethics and Information Technology*, 12: 157-170.

Wendt, Brooke, 2014, *The Allure of the Selfie: Instagram and the New Self-Portrait*, Institute of Network Cultures, Amsterdam.

Wisse, Jakob, 1989, *Ethos and Pathos from Aristotle to Cicero*, Adolf M. Hakkert, Amsterdam.

Neutral phantasies and possible emotions
A phenomenological perspective on aesthetic education

Francesco Pisano

Abstract: In this paper I draw from Husserl's lectures on ethics and manuscripts on phantasy to clarify the role and the structure of aesthetic education within a phenomenological theory of experience. First, I show that Husserl's take on emotions as material contents of value experiences involves the problem of justifying the validity of the relation between factual emotions and ideal values. I then suggest, on the basis of some of Husserl's phenomenological claims on phantasy, that this discrepancy can be bridged through the enjoyment of art: that is, through a process of aesthetic education. I will focus, as Husserl does, on theatre as a case study. His approach to the experience of theatre provides the possibility of an education of emotions by helping the spectator to explore the eidetic structure of emotional states in their individuality, but regardless of their isolated *here and now* (that is, of their facticity). After a presentation of the elements that play into the phenomenological perspective, the first part of the argument refers to the last chapters of Husserl's *Einleitung in die Ethik* (1920/1924). The second part focuses on a 1918 manuscript. I conclude by hinting at the possibility of widening Husserl's account of aesthetic education beyond the experience of theatre.

Keywords: Edmund Husserl; phenomenology; phantasy; emotion; aesthetic education

1. *Introduction*

How does beauty relate to truth? This widely discussed issue recently came back into the spotlight thanks to some studies in cognitive psychology (see, e.g., Schwarz 2018, 25; Reber 2018) that showed how judgments about truth and judgments about beauty share some relevant dynamics concerning their treatment of cognitive information. This topic becomes even more complex when we shift towards the analogous relation between knowledge and art. Since both imply an effort towards a value, philosophers have been discussing their relation from an axiological point of view (see, e. g., Goldie 2008, Sherman and Morrissey 2017). They posed questions such as: could art be useful to knowledge? Can either of them provide some help in the endeavour towards other values?

philinq IX, 1-2021, pp. 29-48
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.269
Submitted: November 2019
Accepted: June 2020

Another complication arises when we focus, in a Platonic fashion, on good as the ultimate aim of beauty and truth – that is, on ethical values. The beauty of art often refers to fictional objects. How can art, often directed towards the production of fictionalities, benefit our attempts at being good people? And also: should it aim to benefit these attempts? This is another often debated topic, nowadays (see, e. g., Carroll 2000, Bermúdez, Gardner 2003).

Some authors (Jenkins 1968, and more recently Spivak 2012: 275-300) claimed that these two clusters of problems (respectively, the relation between art and knowledge and the influence of fiction-based art on ethical endeavours) are in fact connected. They argued that it is only through its fiction-related properties that art can grant some help in the acquisition of a knowledge that, in its turn, helps us to be ethically virtuos. This would be a process of *aesthetic education*.

My aim here is to draw some elements from Husserl's writings in order to sketch the possible articulation of this process from a phenomenological point of view. This would provide a phenomenological framework to experimental studies such as Shoemaker, Costabile & Arkin 2014, which show how our acquaintance with fictional objects or characters does indeed help us better understand and articulate our emotions in view of ethical values. It is worth noting that the phenomenological approach is strikingly absent from the recent reconstruction of the debate about fictional objects in art: for instance, Livingston-Sauchelli (2011) and Brock-Everett (2015) never mention Husserl or other phenomenologists.

Husserl believes that the education we experience through the enjoyment of art is an actual education of real emotions, even if these emotions refer to fictional objects. In the current debate, many would agree with this claim (see, e. g., Gaut 2007: 203-226). A phenomenological perspective, however, offers the specific advantage of allowing a realist approach towards the education of fiction-directed emotions without restricting it to the education of the *behavioural response* to fictional or real objects (as do De Sousa 1990 and Arslan 2014). The phenomenological picture of aesthetic education is neither behaviouristic nor simply reactive. Rather, it has to do both with an active exercise of *phantasy* and with the preparation for *possible* future emotional evaluations and ethical dilemmas.

I will discuss theatre as a case study, since this is Husserl's favourite example. The experience of enjoying a theatre play educates us, the spectators, by supplying us with knowledge about how possible scenarios involving emotions, values and actions could play out. Before discussing this phenomenological perspective, we need to become acquainted with its main ingredients: phantasy as a type of cognitive experience, fiction-directed emotions with their

inherent material logic and art as a privileged field of expression for a specific declination of phantasy.

2. *Phantasy, emotions, aesthetic education*

In a 1918 manuscript,[1] Husserl depicts certain phantasies as *quasi*-experiences (*quasi-Erfahrungen*). What does this mean? Perhaps phantasy is no experience? And, if so, how can it provide any cognitive content whatsoever? Actually, in Husserl's perspective, the *quasi* marks phantasy as a peculiar type of experience, at least from 1918 onwards. And the peculiarity of this experience is the glue that holds the phenomenological picture of aesthetic education together.

In the manuscript, Husserl is mainly concerned with the question of how individuals can appear to an intentional consciousness. First of all, obviously, through perception. Perception shows individuals than are actually there. It shows facts, present here and now. It is our primary access to facticity. However, we can also presentify: through memory, through expectation and through imaginative constructions we can produce or reproduce individual contents for our consciousness. This productive or reproductive domain is the general sphere of phantasy (Husserl 1980: 504-508).

Given this dichotomy between presence and presentification, the manuscript attempts at highlighting the specific primacy of perception among experiences by defining it the only experience that relates to reality (*Wirklichkeit*). It is the only experience that shows a certain *here and now*, a certain facticity (Husserl 1980, 504); thus, it is different from any other kind of experience. However, there are some relevant and widespread experiences that defy this dichotomy, such as primary, immediate retention.[2] The apple we perceived just before the present moment is not, in fact, present here and now. We do not actually perceive it, in a sense, because we can only perceive the apple that is right

[1] The manuscript is presented under the title *Zur Lehre von den Anschauungen und Ihren Modis* (*On the Theory of Intuition and its Modes)* in Husserl 1980: 498-545. It seems that, before 1918, this expression appears only once in Husserl's *corpus*. *Quasi-Erfahrung* cursorily appears in a 1912 manuscript (Husserl 1980: 479) with reference to imagination (*Imagination*, archaic synonym of *Einbildungskraft*, perhaps more focused on *having* the presentified images rather than *producing* them). After 1918, the expression reappears quite often. Husserl focuses again on *quasi-Erfahrungen* in a 1922/23 manuscript, *Reine Möglichkeit und Phantasie* (*Pure Possibility and Phantasy)* in Husserl 1980: 546-570, again with reference to phantasy. *Quasi-Ehrfahrung* even appears in Husserl 1954: 462, with reference to memory. Thus, it is safe to assume that this concept belongs with a certain consistency at least to the later phases of Husserl's phenomenological project.

[2] A comprehensive picture of the relation between primary retention and perception can be found in Hoerl 2013.

here and now. At the same time, it cannot be completely *unwirklich*, irreal, since otherwise no perception of the apple as something that is self-identical through time could be possible. Thus, it has to present itself in some other way. Husserl calls these specific kinds of presentification *perceptual phantasies*, i. e. phantasies that conform to perception, that try to do so, or that pretend to do so. The past aspect of the apple presents itself, Husserl says, as some sort of «suppressed reality» (*aufgehobene Wirklichkeit*: Husserl 1980: 502). It aches to be real, so to speak. This is the meaning of the *quasi*: *quasi*-experience means the *actual* experience of *quasi-perception*, of *as-if perception* (see Bernet 2017). This experience is neutral with respect to facticity (i. e., the positing of real existence) and yet it entertains a certain a priori relation with facts: a relation of conformity. This conformity is simply a consequence of the identity between the perceived individual and the quasi-perceived individual. The apple I just saw and the apple I am seeing now present themselves as the *same* apple.

Let us move on to emotions. In Husserl's late work on ethics, emotions play the role of contents of the will (see Zhang 2009). These contents are, in a few words, everything that has to do with values without being included in a formal axiology. Every will or desire is involved with emotions. According to Husserl, emotions – as opposed to passive feelings (*passive Gefühle*) – are always *emotion-acts* (*Gemütsakte*), active acts of evaluation (Husserl 2004: 3-153). Thanks to these judgment acts, we gain access to the relation between certain values[3] and certain factual situations that are imbued with feelings. Once a value has been found to be possibly related to the situation we are in, we act consequently and in view of this emotionally marked object. These acts are acts of will, since we want to realize the corresponding value. The domain of the acts of will is none other than the domain of ethics (Husserl 2004: 8). Hence, the general form of these acts is a matter of formal axiology. Ethics is, vice versa, a concrete axiology.[4] Emotions are material determinations of the acts of will: they direct them towards a specific value. At the same time, they are factual, contingent moments of these acts. So, the ques-

[3] Concerning the contemplative, autonomous experience of values – such as the grasping of the beauty of something – Husserl seems to hold a realist, perceptualist account. This will not be discussed here, since the main focus is about the value-oriented use of emotions in a practical context, and thus, e. g., the relation between a factual situation and its possible beauty, that has yet to be realized as a new, possible situation. On this issue, see Mulligan 2004.

[4] The critical reception of Husserl's rationalism in ethics presented some good arguments concerning the role of formality and evidence within ethics. The early instances of these objections (presented by Geiger, for instance) directly influenced Husserl's work – prompting him to focus on the role of emotions and feelings of love in the ethical context. Later, the French reception of Husserl (and especially Ricoeur and Levinas) kept on delving into these same issues in an original and radical way. A recapitulation of these objection can be found in Ferrarello 2015: 81-88 and 180.

tion becomes: how can these factual emotional states work as the content a rational law of the will?

Finally, let us consider aesthetic education. The aesthetic sphere could be traditionally described as the sphere of the enjoyment of the work of art. The phenomenological account is characterized by a shift from this classical description to a more dynamic and comprehensive one. The reflection on a work of art is, in Husserl's perspective, a process that trains us to recognize (i. e., to see, to contemplate) possible ethical relations (i. e., practical relations, concerning what we ought to do) through an aesthetic experience (i. e., through an emotional experience under the light of a specific type of phantasy). This process consists of a progressive integration between aesthetics and ethics – that is, of a practical use of disinterested aesthetic contemplation. The possibility of this integration stems from the consideration that our praxis can profitably and repeatedly incorporate an aesthetic moment. In the possibility of this unending refinement lies the opportunity for an exploration of the logic of emotions as structured material contents and material determinations of the will.

The claim that I want to draw from these ingredients is the following: when experienced through art and by virtue of the specific phantastical experience that artistic enjoyment requires, fiction-directed emotions can effectively deepen and articulate our experience of actual emotions. This phantasy is marked by a crucial reference to the reality of our shared world: that is, to the facticity in which our concrete praxis does actually take place. But how would this process articulate itself? To answer this question, we first need a better grasp of the problem of emotions in Husserl's perspective.

3. *The active ethical role of emotions*

Husserl's efforts in defining a phenomenology of practical reason are constantly renewed throughout the course of his philosophical path. Some interpreters (e. g., Ferrarello 2015) argued that ethics are key in understanding Husserl's phenomenological project as a coherent whole. An ethical commitment defines the very character of phenomenology as a theory and as a living praxis.[5] This praxis attempts at extricating a logic of sensibility from experience.[6] Phenom-

[5] The concept of correlation is here intended to embrace also that of coimplication: from an ontological point of view, theory implies (or is encompassed by) praxis, since it *is* a praxis; from an epistemological point of view, instead, praxis implies (or is encompassed by) theory, as long as it is an object of knowledge. About this, see Larrabee 1990.

[6] The Italian reception of Husserl's work put some emphasis on the relation between the phenomenological method and these issues derived from the platonic problem of *methexis* and the Kantian problem of *schematism*. See Paci 1957 and Melandri 1960.

enological ethics could then be designated, within this general framework, as an attempt at describing the correlative logic that embraces both the facticity of emotions and the ideality of values. This correlative and dynamic conception soon ends up being at odds with the notion that the materiality of emotions implies, in a Kantian fashion, their passivity.

At first glance, the juxtaposition of materiality, emotionality and sensibility seems to mark the entirety of Husserl's work. The first instance of a "form-matter approach" about emotions and values dates back to the 1914 lectures about ethics and value-theory (Husserl 1988: 3-153). At least in an initial phase, the idea of an analogy between the value-emotion relation and the form-matter relation grounds Husserl's idea of ethics as a theory of practical reason. The 1920-1924 lectures (Husserl 2004: 3-255), however, put a special emphasis on the material dynamical logic that pertains to emotional states as such.

Husserl recognizes that there is a passive, factual aspect in play within the frame of ethics: it is the aspect of affectivity, i. e. the space of feelings (Husserl 2004: 8). Feelings and affections belong to ethics thanks to their bond with emotions. This bond has an essentially motivational character (see Rump 2017): it is only when we feel something that we are motivated to emotionally evaluate the source of this affection. Hence, emotional acts are both passive (insofar as they are prompted by an affection) and active (insofar as the emotional evaluation involves willing and a degree of self-awareness).

This ambiguous terrain is the field of phenomenological ethics. The form-matter distinction is functional to its clarification. We evaluate if something is to be willed or to be refused by means of the emotional tone of the corresponding experience. This evaluation actively grasps values in their relation with the factual situation we find ourselves in. If we want to consider how these values appear in themselves, we first have to look at what makes each value *a* value *as such* – that is, we need to find the formal laws of the sphere of value (Husserl 1988: 80-101). This formal axiology cannot, however, exhaust ethics. Husserl knows that any ethical theory needs a concrete indication about what we ought to do. Our need to learn what is the right value to prefer in any given emotional situation implies the need to define the proper emotion to cultivate in specific situations. Husserl recognizes this want of a material content for ethical laws even in his 1914 lectures (Husserl 1988: 126-153).

The problem is that we do not just *receive* emotions. They are in fact embedded in a concrete, factual situation. They are interwoven with facticity. In the emotion act we actively extrapolate what relates to a certain value from the variety of what we factually feel. This is the ethical endeavour in its more general form (Husserl 2004, 332). But how is it possible? According to Husserl, the incorporation of feelings within emotions gives a peculiar mo-

tivational, practical and project-oriented light to emotions: which motivate our actions (Husserl 2004: 232-237). This is another way of describing their duplicity: they present themselves within a formal legality, yet they carry on the motivational force received from the feeling embedded in a contingent factual context.

Thus, emotions can work as contents of acts of will (i. e., as material moments of the ethical sphere) only insofar as they are, at the same time, actively motivating an action in view of a value. This evidently clashes with the form-matter framework: not only emotions have their own motivational, material lawfulness (see again Rump 2017), but this lawfulness actually contributes in shaping values by granting a new theoretical and especially *practical* perspective over their relations with different factual situations.

The fact that the emotional access to values is always in view of a possible future action implies that a value has to relate to the unpredictability and contingency of facticity. Thus, our praxis-oriented understanding of values is intimately connected with a perpetual re-establishment of the complex of factual situations the value relates to (See Welton 2000: 309). If we accept a corresponding motivation-oriented conception of emotions, we could say that emotions contribute to our grasping of values to the extent that they help establish the connection between each of these values and a corresponding variety of facts. For all intents, then, the education of emotions is indeed an ethical education.

Once the ethical field has been outlined, the aim of this education becomes to learn how the infinite and nuanced variety of our factual emotional experiences can motivate us towards specific actions and specific values, defined by their position within an eidetic framework. This is difficult, because it deals with the conjuction of two modally heterogeneous extremes: the emotion that is *here* and *now*, radically individual and contingent, and the *ideal* architecture of values. Let us then return to the phenomenological device that makes such an education possible: perceptual phantasy.

4. *Neutrality and phantasy*

Husserl starts his 1904/1905 lectures on *Phantasie und Bildbewusstsein* (*Phantasy and Image-Consciousness*, Husserl 1980: 1-108) by claiming that phantasy is, first of all, a mode of seeing. As said above, phantasy is an actual experience that presents something: only under this light it is possible to describe it phenomenologically (Husserl 1980: 6). Can phantasy present emotions, then? Or, better yet: can it present *examples* of emotions?

The role of the example in phenomenology is well-known.[7] In the first book of *Ideen*, each eidetic intuition is bound to an example, i. e. to an individual whose concrete essence is originally intuitable as an eidetic field of possibilities (Husserl 1976: 14-16). This bond between individual example and eidetic intuition provides a way to think about individuals in an eidetic sense. Each individual implies an individual essence, and this essence is defined by a field of structural relations. An exemplary emotion would then be an individual emotion presented under the light of the material eidetic lawfulness that frames it. This individual exemplary emotion would be somewhat detached from the factual occasion from which it arose, and still be a concrete individual, a motivational force *intuited* as a unity. This is the only way it can remain an emotion (since emotions are innervated by a motivational force, as we have seen) while being at the same time part of a wider eidetic framework as an act of evaluation.

The unity of this exemplary emotion needs a corresponding intuition. Can phantasy provide it? Let us first observe that, if we can produce an image of an object, then this object is at least epistemically possible. As Jansen 2013 claims, images present a certain *situatedness* as long as they involve a spatial and temporal character. Thus, the object of which we produce an image is at least placeable within a possible spatiotemporal frame – that is, within a possible reality. This object is at least a *spatiotemporal* unity. Its actual position is not self-contradictory, even if the object is not part of our shared natural world. A unicorn, for instance, is not impossible, at least from an epistemical point of view. It just did not happen to be any unicorn in the world that we know of.[8]

Now, phantasy can actually provide this kind of spatiotemporal unities without committing to a specific here and now. The key to this resides in its specific *neutrality*. Neutrality is, in a word, the non-positional attitude towards an object, i. e. an assumption of neutrality with respect to the being and non-being of an object. It is a non-positional modification applicable, in principle, to every presentation. Husserl (1976: 250-252) observes that neutrality is not exclusive to phantasy. However, since phantasy is in fact a neutralizing act and

[7] Derrida 1962: 32-37 and 46-47 highlights the relevance of the example as a structural moment of phenomenology as an articulated theory and descriptive method trough some notable observations in his *Introduction* to Husserl's *Origin of Geometry*.

[8] Kripke 1980: 156-158 famously argued against the possibility of unicorns. However, while fictional unicorns and artistic images of unicorns do not seem to involve any consideration regarding their biology or zoology, it seems to me that an actual unicorn would be clearly recognizable through some salient traits. Would a horned horse (with only one horn) suddenly appear within the context of nature, it would be called unicorn. This would perhaps not grant the property of being real to unicorns *according to the classifications of biology and zoology*. However, if this was the case, it seems to me that this should be credited to an insufficiency of these sciences in describing factual reality, rather than to the unreality of the horned horse/unicorn.

is in principle applicable to every presentation in its own right, the two often end up mixed together.[9] An object of neutral cognition is given *as if* it was present in this or that way, while any belief about its real existence, about its facticity, is suspended. Neutral acts are then a species of presentification acts; however, while memory presentifies something that was present then and now is present no more, neutral acts presentify something that was never there in the first place.

Neutral acts entertain a peculiar relation with the facticity we encounter in perceptual experience. They seem to be detached from actual experience and yet related to it. In a neutral cognition we see something that never pertained to factual reality, and yet this cognition provides us with some content about reality. It gives us some information about the eidetic structure of a certain effective material determination considered in its possible variations. It is a modification of reality that also makes us learn something about reality. This double bond of neutral cognition makes it a pivotal resource of phenomenological investigation. However, simply neutralizing an emotional act is not enough: the mere removal of the ontological position can be helpful in a theoretical description of the phenomenological essential texture of a specific emotional state; but it is of no help in exploring the motivational, practical legality of this emotion, since it does not connect it to any possible action nor to any possible factual context.

Phantasy acts are a species of neutral acts. Husserl devotes particular attention to the specificity of phantasy, among other neutral acts, in a manuscript (*Phantasie – Neutralität*, *Phantasy – Neutrality*) wrote between 1921 and 1924 (Husserl 2004: 571-593). Phantasy, he observes, is specifically a disinterested spiritual praxis (Husserl 2004: 577). It does not refer to any present or past stance we could have assumed towards facticity. It is a moment of *Zwecklosigkeit*, of relatively[10] free play, where the *as if* of general neutrality is extended to the operations of the ego itself (Husserl 2004: 572-573). It is a sort of dream-

[9] The idea of neutrality makes its first non-cursory apparition, within the context of Husserl's writings, in a 1912 manuscript (Husserl 1980: 352-364), right before *Ideen*. There, we can find in a footnote (356) the simple equivalence «*neutral = nichtsetzend*» («neutral = non-positive»). But phantasy is not the only non-positioning act. Neutral objects do not need to be necessarily produced or reproduced: we can also neutralise perceived objects. This is the case of the phenomenological *epoché*.

[10] Husserl speaks about a *bound* (*verbunden* or *gebunden*) exercise of phantastical variation. In a formal sense, its bounds are its defining conditions. In a more concrete sense, however, they could perhaps be defined as the ties that keep phantasy within the unity of possible experience, thus giving to phantasy the possibility to present some truth about possible experience as a coherent whole. We could perhaps say that this attention for the dynamic and logically articulated unity of experience is key in characterizing the phenomenological attunement trough which we are trying to reconsider the idea of aesthetic education. On the bounds of phantasy, see Williamson 2016, Summa 2017.

ego,[11] a possible ego as opposed to the actual one, that phantastically delves into an as-if world, into a possible world.

Even as we phantasize, however, we remain the same actual ego. This is the first bond between phantasy and perception. The mark of the as-if does not create a phantastical world, detached from the world of proper (i. e., perceptual) experience. It looks at this same world under a different light. This is why phantasy is an experience from which and to which we can return at will (Husserl 2004: 577). We can bring back some information about our shared reality from this experience, because the neutrality-modification does not necessarily damage the material-eidetic texture of the neutralised object: its sense-contents remain the same, only without facticity.

Moreover, specifically perceptual phantasy can also preserve the spatiotemporal unity of the perceived object. This makes it the much-needed surgical instrument that can remove the specific *hic* et *nunc* correlated with the facticity of an actual emotion while preserving its *hic* et *nunc-ability* – i. e. the spatiotemporal coherency that allows us to imagine that same emotion we are living within other factual contexts. An emotion presents itself and it is actual, here and now. It is caused by this or that. It is a fact. However, if we phantasize about it in a way that conforms to perception, it also becomes *factu-able*, so to say: not only a fact, that maybe now has expired along with its contingent conditions, but *something that could be a fact*, something that could actualize itself in many other possible factual contexts.

As Husserl says in a 1922 manuscript, «the experience in phantasy is possible experience in itself» (Husserl 1980: 548), that goes even beyond the image-consciousness and the imaginary position of *ficta* within the real world. The structure of possible experience, then, is what actual experience and phantasy experience *do* have in common in the most radical sense. Phantasy provides us with modal information about reality: it tells us what is possible and what is not. And it does so a priori, inasmuch as it is not bound by the conditions of a specific *hic et nunc* experience. *Perceptual* phantasy, however, seems to be bound to image-consciousness (Husserl 2004: 504). The question then becomes: does the need to conform to perception necessarily restricts the object of perceptual phantasy to image-objects? And, vice versa: what is the relation between artistic experience and image-objects? Instead of dealing systematically with these issues, Husserl discusses a telling example: that of theatre.

[11] This comparison implies perhaps some problems, given that the very phenomenology of dream experiences is one of the most complex parts of the phenomenological theory of experience. It is, however, a comparison directly suggested by Husserl 1980, 548. For a concise outline of the phenomenology of dreams, see e. g. Zippel 2016.

5. *Perceptual phantasy and aesthetic education: the case of theatre*

Husserl's account of the experience of enjoying a theatre play provides, somewhat between the lines, a way in which emotions can be educated through perceptual phantasy. It is worth noting that, according to Husserl, art is a privileged field for the expression of a specific type of phantasy. It is «the kingdom of phantasy that took form, of perceptual or reproductive phantasy, of phantasy that is intuitive – but also partly not intuitive» (Husserl 1980: 514). This passage of the 1918 manuscript is quite ambiguous. Art is the kingdom of perceptual phantasy or of reproductive phantasy? Or are they the same? Can the conformity to perception only be thought in terms of image reproduction? And what does it mean that a phantasy is intuitive or non-intuitive?

The idea of an affinity between art and neutral cognition is already present in a letter from Husserl to Hugo von Hofmannsthal, dated 1907 (Husserl 1994: 133-136). Even there, theatre is Husserl's preferred example. In the letter, the experience of enjoying a theatre performance is actually compared to the phenomenological *epoché*. The artistic object correlates to a neutralizing act in the 1918 manuscript too. This manuscript also specifies that the general neutrality that the artistic object shares with the pure phenomenon is the neutrality of the objects of intuitive or perceptual phantasy.

This fits particularly well with the features of theatre, even if theatre does not resemble, perhaps, our common figuration of a free play of phantasy. Husserl's focus on this specific form of art mirrors Hume's preference for theatre in his endeavours at describing the relations among passions, values and aesthetic experience – endeavours that provide an account of these relations that is quite similar to the one Husserl presents against Kant (see Hume 1998; Husserl 2004: 200-243).

What is intuitive phantasy, and to what extent does it coincide with perceptual phantasy? We have seen that perceptual phantasy is only partially detached from our actual perceptual world. Since Husserl speaks of the objects of theatre as both object of perceptual and intuitive phantasy, we can refer to them to better understand what Husserl means by the concept of intuitivity. Now, theatre does not necessarily present images that resemble or imitate perceptual reality. However, it does present *ficta*, possibilities that yet are in some way and to some extent informative about our shared actual reality. These ficta conform to the general structure of perception (or to *Perzeption* as opposed to *Wahrnehmung*).[12] The events depicted on stage are indeed fictional: they do not

[12] This distinction is clearly defined in Husserl's MS C16 VI (May 1932). It is however already present in our 1918 manuscript, albeit only in an adjectival form. For simplicity's sake, *Perzeption* is to be understood here as a *Wahrnehmung* without the factual position of its object. It is what makes a

actually happen. What does actually happen is the *acting* of the actors, that we interpret, with the help of perceptual phantasy, as signifying a fictional event. These ficta *do* happen in the as-if world of phantasy (that is: in our shared world posed under the light of the as-if). They are not posed by an ordinary operation of the imagination. As long as we are caught in the suspension of disbelief, with the help of good acting, we react to them as if they were actually happening (Husserl 1980: 514-516). We react through actual feelings and we evaluate the action through actual emotions. The emotions we feel while enjoying a theatre play are facts. But what about the *ficta* to which we relate these emotions? What are they, beyond their fictional character? Are they images?

They are *intuitive unities*. Theatre, as an exercise in collective phantasy, shows us the primary difference between facticity and intuitivity. The phenomenological possibility of a *Perzeption*, of experiencing something *as if* it were perceived, without it actually happening here and now, depends on the possibility of defining an intuitable object without facticity. According to Husserl, this is indeed possible, being that intuitivity (that is, being a possible spatiotemporal individual in *any* perceptual time and space) is not the same as facticity (that is, factually being here and now, in *this* place and time). We can ascribe a specific property – intuitivity – to the phantasy that presents such unities. And intuitivity is essentially what defines perceptual phantasy as such; it is, in fact, the key to its use within aesthetic education.

The enjoyment of a theatre play involving a unicorn requires something more than the simple claim that "here is a unicorn", along with the corresponding phantasy act, does. The unicorn that appears in an artistic experience is indeed a fictional object. It could perhaps be reduced, to some extent, to an image or to a combination of images. But its key peculiarity is that it appears, as a fictional object, neither here nor there, but *in a new possible world* or rather in our shared world under the light of what it could possibly become – a world we enter when we choose to suspend our disbelief and to dissolve the connection of a certain internally coherent experience (a story, an image, and so on) with ordinary practical matters in order to enjoy it aesthetically. This operation of suspension always concerns the whole world, and never a single image; it is a defining moment in the experience of taking part in a theatre play as a spectator (Husserl 1980: 515-517).

The world we take a glance at through a theatre play is indeed a different world, even when overlapped with our shared real world (and practical con-

perceptual phantasy a quasi-perception. The phenomenological possibility of defining the framework of *Perzeption* depends on the possibility of defining a perceptual object with a character other than facticity. According to Husserl, this is indeed possible, being that intuitivity (that is, being a spatiotemporal individual) is not the same as facticity (that is, factually being here and now).

text). We generally do not believe that the Venice of Shakespeare's *Othello* is the same as the real one, or that the killing of Desdemona happened there and then in this exact way, or even that it is happening right at this moment on the stage. But it is indeed a possible world, that remains connected to our actual world through the same relation that ties together facts and possibilities. The possible actions that constitute the narration are not isolated images in search of a spatiotemporal collocation: rather, they already are defined individual (possible) facts within a possible world, with their own spatiotemporal position. This is why theatrical *ficta* cannot be images in the sense in which an image is a refiguration-of (*Abbild von*). How can there be a refiguration of a world, since we can never figure a world as a whole in the first place? The actors on stage do not portray an imitation of reality: they transport us in an artistic illusion (*künstlerische Illusion*) that they prompt our perceptual phantasy to produce (Husserl 1980: 515-516).

How does this happen? We, not being able to reproduce a full image of a world, simply borrow the structure of a world from perceptual reality. Each fictional event takes place in the spatiotemporal frame that is the world of perception, unless stated otherwise within the play. And, even then, the framework to which we apply any suggested spatiotemporal modification is the one that we borrow from perception. Dramatic action takes place within the space of this possible world, thus requiring that this world be defined by certain conditions that this fictional world shares with the real and ordinary context of our actions. This framework provides some unity conditions for the events taking place in it: this happens here, then that happens there, and so on.

One of phenomenology's revolutionary concepts is indeed that of intuition, or of direct grasping of non-perceptual unities (Hintikka 2003). Thus, *this* and *that* are indeed objects of an intuitive phantasy, since they are non-perceptual spatiotemporal unities. By claiming that the non-intuitive aspect of artistic phantasy is the pivotal ingredient of this experience, since the time and the space of the *ficta* are not completely adherent to the time and space of perception, Carreño 2016 neglects this basic form of spatiotemporal unity: space and time can indeed vary their structure within fiction; but they can do so only insofar this basic form of *intuitive* individuation first allows this fiction to be distinguished from complete chaos.

Perceptual phantasy is indeed still phantasy, inasmuch as it represents a (relatively) free play of as-if possibilities within certain defining conditions, and does so a priori, independent of the specific occasionality of each actual experience. But it also the pivotal type of intuitive phantasy. This allows for a structural convertibility of phantasy individual with real ones. And this is the key of perceptual phantasy's possible use within ethics.

Let us take a closer look at this conversion between fictional facts and *hic et nunc* facts. The 1918 manuscript starts with some considerations about the occasionality of perceptual content. The perceptual object is given in such a way that its presentation is bound to an irreducible, non-repeatable and non-describable spatiotemporal contingency (Husserl 1980: 499). This contingency belongs to its concrete individual essence: it cannot be completely described, but it can be originally given in a perception. A perceptual phantasy seems a *contradictio in adjecto* only if we assume that this original givenness, this intuitivity, is one and the same with the positionality we attribute to every natural perception – the same positionality every neutral act renounced to. If intuitivity is *not* positionality, then individuality and occasionality are not the same, and it is possible to distinguish a general perceptual mark from each concrete *Wahrnehmung*. This mark (the mark of intuitivity and spatiotemporal individuality) is the abstract structure that Husserl designates as *Perzeption*. How can artistic phantasy conform to *Perzeption* while meeting the criteria that define it as a phantasy? It has to remain intuitive, i. e. referred to individuals.

We could say that perceptual phantasy is an intuitive phantasy as long as its objects are marked by a character of spatiotemporal individuation. They do not necessarily need to be reproductive figurations of reality. Theatre performances often present objects that are most evidently convertible in real, possible situations, since we are used to unify the individuality of real objects with other characteristics such as causal interactivity, time linearity and space continuity. However, individuality – the objective correlate of intuitivity – does not necessarily need to be *factual* individuality.

Intuitivity, as Husserl puts it, means the structurally correlative identity between a possible object of phantasy and a possible object of perception. Perceptual phantasy does, in fact, grasp individuals: it grasps free possibilities that are nevertheless marked, in some way, by a character of spatiotemporal individuation. This is proven by the fact that the phantastical object *A* (the individual-character) and the actual object *B* (the individual-actor) phenomenologically converge – as Husserl writes (Husserl 1980: 508) – in their individual essences, and yet they are different in that the phantastical singularity is modified by the mark of phantasy. They are suspended in a possible, but not yet accomplished, identity. The phenomenology of perceptual phantasy in an artistic experience is the phenomenology of this suspended contrast, or suspended convergence. The possibility of this identification suffices in defining the phantastical object as an individual, since 1) *A* has to be susceptible of being identified with B; 2) B is determined as an individual; 3) the only difference between A and B follows from a neutrality modification that, as we have seen, has to do only with the position of the object and not with its determination-content (or sense-

content).[13] Thus, A – the fictional object of artistic and perceptual phantasy – is also an individual in itself, and perceptual phantasy is inherently intuitive while remaining neutral, i. e. not positional.

This convertibility between individuals is what allows us to be emotionally and rationally involved in a play, repeating, evaluating and enriching the choices prescribed by the author (Husserl 1980: 519). In doing so, we explore how different emotions can act as different motivations, and how different motivations could prompt different and unexpected consequences. As Husserl claims, our evaluations concerning the action happening in the as-if world «possess a sort of objective truth, even if they are about ficta […]. Indeed we, as actual men, judge, and not the poet in a predetermined way» (Husserl 1980: 520). This openness of the work of art, this request to judge that the work of art poses to us, is the essential prompt of aesthetic education.

Elicited by this request, we explore through fiction-directed emotions the logical articulation of possible emotional states. Eventually, we can draw an eidetic cartography of the ethical sphere, encompassing possible emotions, possible desires and their internal value-structure. We can understand, for instance, that an excess of pleasure structurally converts into a certain type of pain, and from this eidetic structure we could extract the rational norm that commands moderation in the experience of pleasure, given that we partake to the rational pursuit of happiness. Classic examples of this exercise are offered by Shakespeare's *Macbeth* (through which we investigate personal pride as a motivation for political ambition) and Sophocles' Theban plays (through which we delve into the reciprocal implications of rage, sorrow and revenge).

6. *Beyond theatre?*

Each one of the possibilities grasped by perceptual phantasy is an individual: a *this*, an *individuum*. And yet, it is a *possible* individuum, in relation to other eidetic material possibilities (or impossibilities). We have seen that this intuitive unity is not an image. The last question I will consider here is if this means that even less perception-bound forms of art can present this kind of exemplary individuals.

The idea of a free-playing phantasy would seem to concern an abstract painting more than theatre.[14] Literature and music also help us phantastically

[13] The identification of the determination content of an object with its (noematic) sense remains an open and problematic possibility. Here, I refer especially to Husserl 1976: 297-299.

[14] Abstract paintings represent an interesting borderline case for the "narrative" conception of art that Husserl emphasizes here. It could be argued that even abstract paintings tell us about possible courses of action, at least in a wide and perhaps metaphorical sense. Abstract paintings do in fact

exploring and articulating our emotions. If we think about works such as Melville's *Moby Dick* and Mozart's *Requiems* we can perhaps figure out how these forms of expression can help us better understand pride, rage and sorrow. But is the phantasy involved in music and literature intuitive, at least to a certain extent?

We have seen that it is not contradictory to attribute a non-figurative intuitivity to music, literature or even to abstract visual art, since intuitivity and figurative character are different properties. The problem is to find these intuitive unities in other artistic experiences. In some arts, they are somewhat circumscribable: *this* is what the painting represents, *this* is the event the actors are acting. Now, the intuitive character of these unities does not reside in these visual or verbal explanations. It is in the "this", since intuitive unity is exactly what allows to speak of *this* and *that* fictum in the first place. And it is easy to recognize that even an abstract painting can express *a* "this", because otherwise we could never be able answer to questions such as "what does it talk about?" or "what does it express?". This impossibility would be structural. Instead, it seems that our possible inability to answers depends on the fact that we have not reflected enough on the artwork, while the artwork in itself does indeed express *something*.

Here I want to just cursively note that a phenomenological theory of fictional deixis does exist. A successful integration of a phenomenological theory of fictional deixis (see the concept of deixis *am Phantasma* in Bühler 1934: 124) within Husserl's framework would allow to complete our picture of a phenomenological aesthetic education. This integration would perhaps allow to speak of the contents of art in general as intuitive individual essences of "spiritual", intersubjective, non-sensible relations. They would be intentional objects connotated with certain emotions and referring to certain values. Picasso's *Guernica* would connote the "values" of violence and war with the emotions of pain and sorrow. Malevich's *White on white* would connect a feeling of mystic abandonment with the values of simplicity and absoluteness. In any case, these relations can be simply considered the (metaphorical) meaning of the abstract artistic product, which works as a sort of signifier or as a form of expression.[15] All these instances do in fact provide some sort of objective information about our shared

consists of perceptual elements installed in certain interrelations. If we could present this interaction of perceptual elements as the exploration of possible interactions between emotions (and thus values), then the account of artistic experience that we are exploring could also encompass abstract art. For a phenomenological introduction to the issue, see Crowther 2009: 99-119. The same problem arises when dealing with non-tonal music and, in general, with all art forms that seem to refuse any conformity to perception.

[15] On the problem of defining the object of abstract artistic expression, see Poggi 2004.

factual reality, as long as these fictional emotions are facts on their own and can be converted into possible non-fictional emotions (Husserl 1980: 520).

In short, perceptual, intuitive phantasy entertains a peculiar relation with truth, since it can show possible relations between certain values and certain factual situations. Its neutrality dissolves the occasional constraints of actual perception. Its phantastical character allows for a relatively free exploration of eidetic possibilities. Its intuitivity allows this eidetic exploration to refer to individual actions and desires. We, as spectators, reflect on the individual fiction-directed emotions we feel during the fruition of art – on the emotional acts through which we connected certain situations, certain factual feelings, with certain values. As we partake in this phantastical simulation, we evaluate and phantasize about other possible emotional evaluations and about other possible actions. We continue the phantasy that constitutes the artistic object. We explore it, we articulate it, and we detach from the object and return to it. And, insofar as this eidetic exploration also concerns examples of emotions, we gain a better understanding of our own emotionality, and thus of our own ethical disposition.

This is indeed a form of aesthetic education, albeit a peculiar one. It searches no more, within the horizon of beauty, an ideal medium between Kantian opposites. It ends up putting into question the very distinction between form and matter. The highlight that a phenomenological approach puts on the importance of a possible practical use of phantasy experiences shows that there is just as much ideality and formal lawfulness within the factual space of emotions, as there is facticity and material specificity within the ideal space of values. A logic that is both relatively material and relatively formal already can embrace the entirety of experience, i. e. both fictional and non-fictional objects.

7. *Conclusion*

Kind (2016) treats imagination, when engaged with art, as unconstrained. This paper showed that even when we zoom out from imagination to non-imaginative phantasy some constraints remain for the phantastical involvement in an artistic experience. It did so by sketching the structure of a logico-phenomenological foundation of the idea of aesthetic education. It also clarified Husserl's idea of a connection between aesthetics and ethics by explicating and presenting what was already implicit in his phenomenological work.

It argued that we can train ourselves to progressively understand our actual emotions, however occasional, as motivations towards a certain value and towards a certain action by repeatedly and freely comparing them to exemplary fiction-directed emotions. The individuality of these exemplary emotions

does not prevent their understanding as material (and eidetically structured) contents of a rational law of the will. Hence there is no real contrast, in Husserl's lifelong work, between a Kantian (Crespo 2015) and an anti-Kantian or Aristotelic (Drummond 2014) perspective on ethics. Even the juxtaposition of the two different perspectives along two different phases of Husserl's work on ethics (Smith 2007) does not make proper justice to the profound unity of Husserl's phenomenological take on ethics.

The phenomenological idea of aesthetic education presented here could contribute to the effort in highlighting the intrinsic ideal relational logic of our living experience against unilaterally "naturalizing" approaches. It could also provide an argument for the ethical value of art that does not commit to any behaviouristic verification and presents itself with clear evidence within self-reflection and dialogical confrontation. A more complete picture of the promising phenomenological possibilities in the field of aesthetic education requires, however, further investigations on the relations between non-figurative art and intuitive phantasy.

Francesco Pisano
francesco.pisano@unifi.it
University of Florence

References

Arslan, Asuman Aypek, 2014, "A study into the effects of art education on children at the socialisation process", in *Social and Behavioural Sciences*, 116: 4114-4118.

Bermúdez, José Luiz, Gardner, Sebastian, 2003, *Art and Morality*, Routledge, London.

Bernet, Rudolf, 2017, *A Husserlian Analysis of Imagining What is Unreal, Quasi-Real, Possibly Real, and Irreal*, <https://cfs.ku.dk/summer-school-2017/background-readings/phantasy-0117-final.pdf> (15/10/2020).

Brock, Stuart, Everett, Anthony, 2015, *Fictional Objects*, Oxford University Press, Oxford.

Bühler, Karl, 1934, *Sprachtheorie*, Fischer, Jena.

Carreño, Javier Enrique, 2016, "Husserl on the Right Timing of Depictions", in Benedek András, Veszelszki Ágnes, eds., *In the Beginning was the Image: The Omnipresence of Pictures*, Peter Lang, Frankfurt am Main: 59-66.

Carroll, Noël, 2000, "Art and Ethical Criticism: An Overview of Recent Directions of Research", in *The Journal of Aesthetic Education*, 2 (3): 21-39.

Crespo, Mariano, 2015, "Husserl on Personal Aspects of Moral Normativity", in *Ethical Perspectives*, 22 (4): 699-722.

Crowther, Paul, 2009, *Phenomenology of the visual arts (even the frame)*, Stanford University Press, Stanford.

De Sousa, Ronald, 1990, "Emotions, Education and Time", in *Metaphilosophy*, 21 (4): 434-446.

Derrida, Jacques, 1962, *Introduction à l'Origine de la Géométrie de Husserl*, Presses Universitaires de France, Paris.

Drummond, John, 2014, "Husserl's Phenomenological Axiology and Aristotelian Virtue Ethics", in Miira Tuominen, Sara Heinämaa, Mäkinen Virpi, eds., *New Perspectives on Aristotelianism and Its Critics*, Brill, Leiden: 179-195.

Ferrarello, Susi, 2015, *Husserl's Ethics and Practical Intentionality*, Bloomsbury, London.

Goldie, Peter, 2010, "Virtues of Art", in *Philosophy Compass*, 5 (10): 830-839.

Hintikka, Jaakko, 2003, "The Notion of Intuition in Husserl", in *Revue Internationale de Philosophie*, 2003/2 (224): 57-79.

Hoerl, Christoph, 2013, "Husserl, the Absolute Flow, and Temporal Experience", in *Philosophy and Phenomenological Research*, 86 (2): 376-411.

Hume, David, 1998, *Of Tragedy*, Hume, David, *Selected essays*, Oxford University Press, Oxford-New York: 126-133.

Husserl, Edmund, 1976, *Ideen Zu Einer Reinen Phänomenologie Und Phänomenologischen Philosophie. Allgemeine Einführung in Die Reine Phänomenologie*. Edmund Husserl Gesammelte Werke (Husserliana), Bd. III/1, Nijhoff: Den Haag.

Husserl, Edmund, 1980, *Phantasie, Bildbewusstsein, Erinnerung. Zur Phänomenologie der anschaulichen Vergegenwärtigungen. Texte aus dem Nachlass (1898-1925)*, Edmund Husserl Gesammelte Werke (Husserliana), Bd. XXIII, Nijhoff: Den Haag.

Husserl, Edmund, 1988, *Vorlesungen über Ethik und Wertlehre 1908-1914*, Edmund Husserl Gesammelte Werke (Husserliana), Bd. XXVIII, Kluwer, Dordrecht-Boston-London.

Husserl, Edmund, 1994, *Briefwechsel. Band VII. Wissenschaftlerkorrespondenz*. Edmund Husserl Gesammelte Werke (Husserliana), Dokumente, Bd. III, Kluwer, Dordrecht-Boston-London.

Husserl, Edmund, 2004, *Einleitung in die Ethik. Vorlesungen Sommersemester 1920/1924*, Edmund Husserl Gesammelte Werke (Husserliana), Bd. XXXVII, Springer: Dordrecht.

Jansen, Julia, 2013, "Imagination, Embodiment and Situatedness: Using Husserl to Dispel (Some) Notions of 'Off-Line Thinking'", in Rasmus Thybo Jensen, Dermot Moran, eds., *The Phenomenology of Embodied Subjectivity*, Springer, Basel: 63-79.

Jenkins, Iredell, 1968, "Aesthetic Education and Moral Refinement", in *The Journal of Aesthetic Education*, 2 (3): 21-39.

Kind, Amy, 2016, "Imagining under Constraints", in Amy Kind, Peter Kung, eds., *Knowledge through Imagination*, Oxford University Press, Oxford: 145-159.

Larrabee, Mary Jeanne, 1990, "The Contexts of Phenomenology as Theory", in *Human Studies*, 13/3: 195-208.

Livingston, Paisley, Sauchelli, Andrea, 2011, "Philosophical perspectives on fictional

characters", in *New Literary History: A Journal of Theory and Interpretation*, 42 (2): 337-360.

Melandri, Enzo, 1960, *Logica ed esperienza in Husserl*, Il Mulino, Bologna.

Mulligan, Kevin, 2004, "Husserl on the 'Logics' of Valuing, Values and Norms", in Beatrice Centi, Gianna Gigliotti, eds., *Fenomenologia della Ragion Pratica: L'Etica di Edmund Husserl*, Bibliopolis, Napoli: 177-225.

Paci, Enzo, 1957, "Relazionismo e schematismo trascendentale", in Id., *Dall'esistenzialismo al relazionismo*, D'Anna, Firenze: 175-238.

Poggi, Stefano, 2014, *L'anima e il cristallo. Alle radici dell'arte astratta*, Bologna, Il Mulino.

Reber, Rolf, 2018, "Beauty and Truth in Mathematics. Evidence from Cognitive Psychology", in Sorin Bangu, ed., *Naturalizing logico-mathematical Knowledge. Approaches from Philosophy, Psychology and Cognitive Science*, Routledge, New York: 252-267.

Rump, Jacob Martin, 2017, "The Epistemic Import of Affectivity: A Husserlian Account", in *Midwest Studies in Philosophy*, 41 (1): 82-104.

Schwarz, Norbert, 2018, "Of fluency, beauty and truth: Inferences from metacognitive experiences", in Joëlle Proust, Martin Fortier, eds., *Metacognitive Diversity. An Interdisciplinary Approach*, Oxford University Press, Oxford: 25-46.

Sherman, Aleksandra, Morrissey, Clair, 2017, *What Is Art Good For? The Socio-Epistemic Value of Art*, viewed on 05/10/2020, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5581397/>.

Shoemaker, Randi, Costabile, Kristi, Arkin Robert, 2014, "Self-Expansion through Fictional Characters", in *Self & Identity*, 13 (5): 556-578.

Smith, David Woodruff, 2007, *Husserl*, Routledge, London.

Summa, Michela, 2017, "Experiencing Reality and Fiction. Discontinuity and Permeability", in Michela Summa, Thomas Fuchs, Luca Vanzago , eds., *Imagination and Social Perspectives. Approaches from Phenomenology and Psychopathology*, Routledge, New York: 45-64.

Welton, Donn, 2000, "Husserl and the Japanese", in Welton, Donn, *The Other Husserl: The Horizons of Transcendental Phenomenology*, Indiana University Press, Bloomington: 306-327.

Williamson, Timothy, 2006, "Knowing by Imagining", in Amy Kind, Peter Kung, eds., *Knowledge through Imagination*, Oxford University Press, Oxford: 113-123.

Zhang, Wei, Xin Yu, 2009, "The Foundation of Phenomenological Ethics: Intentional Feelings", in *Frontiers of Philosophy in China*, 4(1): 130-142.

Zippel, Nicola, 2006, "Dreaming Consciousness: A Contribution from Phenomenology", in *Rivista Internazionale di Filosofia e Psicologia*, 7(2): 180-201.

Blameworthiness, Willings, and Practical Decisions

E.J. Coffman

Abstract: What kinds of things can we be morally responsible for? Andrew Khoury offers an answer that includes (i) an argument for the impossibility of blameworthiness for overt action, and (ii) the assertion that "willings are the proper object of responsibility in the context of action". After presenting an argument for the inconsistency of Khoury's answer to our focal question, I defend the following partial answer that resembles, but differs importantly from, Khoury's answer: one can be blameworthy for a practical decision – that is, an essentially intentional momentary mental action of forming an intention to do something that resolves prior felt unsettledness about what to do.

Keywords: moral responsibility; blameworthiness; resultant moral luck; action; willing; practical decision

What kinds of things can we be morally responsible for? Andrew Khoury (2018: 1368) offers an answer that includes (i) an argument for the impossibility of (non-derivative[1]) blameworthiness for overt action (that is, action essentially involving peripheral bodily motion), and (ii) the assertion that "willings are the proper object of responsibility in the context of action". Unfortunately, Khoury's answer to our focal question is inconsistent: if his argument for the impossibility of blameworthiness for overt action is sound, then so is a similar argument for the impossibility of blameworthiness for what he calls 'willings'. After presenting an argument for the indicated conditional claim, I defend the following partial answer to our focal question that resembles, but differs importantly from, Khoury's answer: one can be (non-derivatively) blameworthy

[1] Non-derivative blameworthiness for X is blameworthiness for X that does not derive from blameworthiness for something other than X. Derivative blameworthiness for X is blameworthiness for X that derives from blameworthiness for something other than X. Khoury (2018: 1363) rejects "the possibility that a person could be blameworthy for [X] even though she is blameworthy [for X] on the basis of something else." This commits Khoury to the thesis that all blameworthiness is non-derivative blameworthiness (cf. Khoury 2018: 1375). Henceforth, 'blameworthiness' abbreviates 'non-derivative blameworthiness'.

philinq IX, 1-2021, pp. 49-56
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.293
Submitted: May 2020
Accepted: September 2020

for a practical decision (cf. Pereboom 2001) – that is, an essentially intentional momentary mental action of forming an intention to do something that resolves prior felt unsettledness about what to do.[2]

To understand Khoury's answer to our focal question, we must understand what 'willing' means in this context. In the following passages, Khoury partially clarifies what he means by 'willing':

> …[W]illings are mental events, expressive of agency, that occur when we act. They are the necessary mental component of action that accounts for the voluntary nature of such events… Willings are what is left of an action after we strip away all of its contingent results. (1364)

> The willing is… the "actish" event that occurs when a properly envatted agent reasonably believes that she is acting… [T]he existence of willings… should be no more controversial than the existence of actions. (1364n19)

> [T]he mental component of action, what I have called the willing, is metaphysically separable from the bodily movement and its further consequences, in the sense that it could occur in the absence of the bodily movement or its consequences… On physicalism this event will entail some physical events, presumably in the brain… Whenever we act there is an associated mental event, what I have called a willing… (1365)

> The expression of agential control that is necessary for action is what I have been calling a willing… [I]t is only through willings that we are able to interact with each other as agents. (1366-1367)

We can further clarify what 'willing' means in this context by considering the following list of action-related items, each of which is such that its "existence is no more controversial than the existence of actions": intentions; active intention-acquisitions (= practical decisions); passive intention-acquisitions; attempts to (perform an instance of action-type) A; beginning portions of attempts to A.[3]

Willings aren't intentions, for willings are events but intentions (like desires and beliefs) are states. Willings aren't *active* intention-acquisitions, for we can act without actively acquiring an intention (that is, without making a practical decision). Willings aren't *passive* intention-acquisitions, for passive intention-acquisitions aren't expressions of *agential* control. Finally, willings

[2] For helpful discussion of practical decision, see chapter 2 of Mele 2017.

[3] This paragraph and the next one are indebted to Adams and Mele (1992) as well as Clarke and Reed (2015).

aren't attempts to act, for many attempts to act involve peripheral bodily motion. Hence, by 'willing', Khoury must mean *the beginning portion of an attempt to act*. Beginning portions of attempts to act – for short, 'attempt initiations' – are brain events that (i) result from acquisitions of proximal intentions (that is, intentions to perform an instance of a certain action-type straightaway); (ii) express or manifest agential control; and (iii) happen whenever we act. Moreover, a typical attempt initiation will be accompanied by a reasonable belief that the pertinent agent is acting, for the initiation of an attempt to A will typically produce an experience as of A-ing, which experience constitutes evidence that its subject is A-ing.

Having further clarified what Khoury means by 'willing', we can now turn to his argument for the impossibility of blameworthiness for overt action, which he helpfully summarizes in the following passage:[4]

> For any bit of behavior that any agent engages in, there will be a hypothetical mental twin who has been recently envatted. And, insofar as we deny resultant moral luck, we should also deny that there is a difference in [blameworthiness] between the normal agent and the envatted mental twin. If so, then the possible objects of [blameworthiness] are limited to elements that remain fixed across such cases: elements of the agent's mental life. (1361)

Here's a more general and formal statement of Khoury's (2018: 1358-1363) argument for the impossibility of blameworthiness for overt action:

> (1) Necessarily, if one is (non-derivatively) blameworthy for an overt action A, then one is more blameworthy than one would have been had one merely tried unsuccessfully to A.
> (2) Necessarily, if one is blameworthy for an overt action A, then one is not more blameworthy than one would have been had one merely tried unsuccessfully to A.

Therefore,

> (3) Necessarily, if one is blameworthy for an overt action A, then one both is *and* isn't more blameworthy than one would have been had one merely tried unsuccessfully to A. [1,2]
> (4) Necessarily, it's false that one both is *and* isn't more blameworthy than one would have been had one merely tried unsuccessfully to A.

[4] Notably, each of the three cases that Khoury (2018: 1358-1360) presents and discusses prior to providing the summary statement of his argument focuses exclusively on blameworthiness.

Therefore,

(5) Necessarily, it's false that one is blameworthy for an overt action A. [3,4]

What should we make of this argument?

I agree with Khoury that the (1)-(5) argument is sound.[5] But if the (1)-(5) argument is sound, then so is the following similar argument:

(6) Necessarily, if one is (non-derivatively) blameworthy for initiating an attempt (or, for starting to try) to A, then one is more blameworthy than one would have been had one merely possessed an ineffective proximal intention to A (that is, an intention to A straightaway that never actually initiates an attempt to A).[6]

(7) Necessarily, if one is blameworthy for initiating an attempt to A, then one is not more blameworthy than one would have been had one merely possessed an ineffective proximal intention to A.

Therefore,

(8) Necessarily, if one is blameworthy for initiating an attempt to A, then one both is *and* isn't more blameworthy than one would have been had one merely possessed an ineffective proximal intention to A. [6,7]

(9) Necessarily, it's false that one both is *and* isn't more blameworthy than one would have been had one merely possessed an ineffective proximal intention to A.

Therefore,

(10) Necessarily, it's false that one is blameworthy for initiating an attempt to perform A. [8,9]

The soundness of the (1)-(5) argument entails the soundness of the (6)-(10) argument provided both that (1) entails (6) and that (2) entails (7). I'll now argue for each of these entailment claims. First: If (1) is true, then (6) is as well.

[5] Many other theorists would join Khoury in deeming the (1)-(5) argument sound, including Davison (1999), Zimmerman (2002, 2006), and Graham (2017).

[6] See Graham (2017: 171-172) and Mele (2017: 35ff.) for cases involving agents with ineffective proximal intentions.

Here's the general principle in virtue of which (1) is true:

> (GP) If one is (non-derivatively) blameworthy for the occurrence of an event E, then one is more blameworthy than one would have been had E not occurred.

Observe that (GP) also entails (6). Hence, if (1) is true, then (6) is as well.

Second: If (2) is true, then (7) is as well. (2) is extremely plausible in light of reflection on a pair of cases each of which involves an agent who attempts to commit murder (cf. Khoury 2018: 1358-1359). In the first case, the murder attempt succeeds. In the second case, the murder attempt fails, but only because (say) an unfortunate bird flies into the path of a bullet. Intuitively, the would-be murderer is no less blameworthy than is the actual murderer. This intuition about the pertinent pair of cases justifies (2). Turning now to (7), consider a pair of cases each of which involves an agent who acquires a proximal intention to commit murder (that is, an intention to commit murder straightaway). In the first case, the agent's acquisition of the relevant proximal intention immediately causes the beginning portion of a murder attempt.[7] In the second case, the agent's acquisition of the proximal intention doesn't cause the beginning portion of a murder attempt, but only because (say) the agent's brain has been surreptitiously altered by a skilled neurosurgeon. Intuitively, the agent with the ineffective intention (for short, the 'ineffective intender') is no less blameworthy than is the agent with the effective intention (for short, the 'effective intender').[8] Moreover, the claim that the ineffective intender is as blameworthy as is the effective intender is itself no less plausible than is the earlier comparative claim that supports (2). Accordingly, if (2) is justified by reflection on the pair of cases involving the would-be and actual murderers, then (7) is justified by reflection on the pair of cases involving the ineffective and effective intenders. If (2) is true, then (7) is as well.

I conclude, then, that the (6)-(10) argument is sound if the (1)-(5) argument is sound. In short, a (non-derivatively) blameworthy attempt initiation would be just as much an instance of resultant moral bad luck as would be a blameworthy overt action; and so, since resultant moral bad luck is impossible (cf. Khoury 2018: 1369-1375), there can't be a blameworthy attempt initiation. Khoury's answer to the question what we can be blameworthy for is therefore

[7] Cf. Mele (1992: 167): "[T]he mental and physical architecture of any being capable of intentional action is such that when such a being acquires a proximal intention to A, an immediate effect is the triggering of appropriate actional mechanisms, unless something prevents this."

[8] Cf. Graham (2017: 169): "All the blame that is appropriately borne toward someone in response to her performing some action is exhausted by the blame appropriate in response to her having the intention she has to perform that action in the situation" (see also Zimmerman 2002, 2006).

inconsistent: if his argument for the impossibility of blameworthiness for overt action is sound, then it's false that "willings are the proper object of responsibility in the context of action" (1368).

In the balance of this note, I'll highlight and defend a partial answer to our focal question that resembles, but differs importantly from, Khoury's answer. It is this: one can be (non-derivatively) blameworthy for a practical decision (cf. Pereboom 2001) – that is, an essentially intentional momentary mental action of forming an intention to do something that resolves prior felt unsettledness about what to do. Unlike Khoury's claim that we can be blameworthy for attempt initiations, the thesis that we can be blameworthy for practical decisions is consistent with the soundness of the (1)-(5) and (6)-(10) arguments.

Consider the following decision-focused analogue of the (1)-(5) and (6)-(10) arguments:

> (11) Necessarily, if one is (non-derivatively) blameworthy for actively forming a particular intention (that is, for making a particular practical decision), then one is more blameworthy than one would have been had one not actively formed that intention.
>
> (12) Necessarily, if one is blameworthy for actively forming a particular intention, then one is not more blameworthy than one would have been had one not actively formed that intention.[9]

Therefore,

> (13) Necessarily, if one is blameworthy for actively forming a particular intention, then one both is *and* isn't more blameworthy than one would have been had one not actively formed that intention. [11,12]

[9] Writes Michael Zimmerman (2006: 605): "…[T]he fortuitous intervention of nature in the form of a passing bird, while reducing the scope of [an actually successful] assassin's culpability, would not diminish its degree. But… nature could intervene earlier in the sequence of events from [the assassin's decision to the flight of the bullet]; indeed, it could intervene even *prior* to… the assassin's decision to shoot. For example, it could happen that, just as he is about to make this decision, the assassin is seized by a sudden sneeze that prevents him from making it. If the fortuitous intervention of the bird does not diminish his culpability, I cannot see how the fortuitous intervention of the sneeze could do so." These remarks would seem to commit Zimmerman to (12). In any case, the thesis labeled '(16)' in the text (see below) enables the following explanation of how the sneeze could diminish the assassin's culpability: due to the sneeze, the assassin remains unsettled about whether to shoot the potential victim, and as yet might not actively settle on doing so; accordingly, the assassin isn't yet as blameworthy as he would be had he already actively settled on shooting the potential victim.

(14) Necessarily, it's false that one both is *and* isn't more blameworthy than one would have been had one not actively formed a particular intention.

Therefore,

(15) Necessarily, it's false that one is blameworthy for actively forming a particular intention. [13,14]

What should we make of this argument?

While clearly similar to the successful (1)-(5) and (6)-(10) arguments above, the (11)-(15) argument fails due (12)'s implausibility. To begin to see this, note that (12) lacks the intuitive plausibility of (2) and (7), the corresponding steps in (respectively) the (1)-(5) and (6)-(10) arguments: a practical decision (= an essentially intentional active intention formation that resolves prior felt unsettledness about what to do) is a better candidate for thing that can increase one's degree of blameworthiness than is either an overt action or an attempt initiation.[10] Moreover, (12) should strike one as implausible in light of reflection on the following thesis:

(16) One's actively settling upon executing a particular action-plan (say, to kill a sworn enemy) could make one at least a *bit* more blameworthy than one would be were one still unsettled about whether to execute the relevant action-plan.

(16) obviously entails the denial of (12). So, to the extent that one finds (16) plausible, one should find (12) implausible. But (16) is extremely plausible (cf. Davison 1999: 248-9). Hence, we should deem (12) implausible. I conclude that we can see our way past the (11)-(15) argument by first contrasting (12) with (2) and (7), and then reflecting on (16) and its bearing on (12).

Finally, reflection on (16) also yields an error theory for the following claim that Khoury (2018: 1368) makes on behalf of the assertion that "willings are the proper object of responsibility in the context of action" (1368):

[10] According to Pamela Hieronymi (2006: 56), "[i]t is now quite standard… to think of intending as settling the question of what one will do. Having settled that question… leaves one open to certain questions and criticisms… Thus an intention… seems at least in part commitment-constituted. An intention is a commitment to doing something." Anyone who accepts this account of intentions should find appealing (the thesis expressed by) the sentence to which this note is appended as well as the thesis labeled '(16)' in the text.

…[T]he fact that the strength of one's desires or motivations has crossed the volitional threshold to produce a willing in a particular context has direct evaluative significance.

While the (6)-(10) argument impugns Khoury's claim here, his claim is quite similar to – and therefore easily conflated with – the extremely plausible (16). Khoury's claim is quite similar to (16) because attempt initiations are quite similar to practical decisions: attempt initiations, like practical decisions, are essentially intentional momentary actions that happen in the brain (cf. Clarke and Reed 2015: 7-12).

E.J. Coffman
ecoffma1@utk.edu
University of Tennessee, Knoxville

References

Adams, Frederick, and Mele, Alfred R., 1992, "The Intention/Volition Debate", in *Canadian Journal of Philosophy*, 22: 323-338.

Clarke, Randolph, and Reed, Thomas, 2015, "Free Will and Agential Powers", in *Oxford Studies in Agency and Responsibility*, 3: 6-33.

Davison, Scott A., 1999, "Moral Luck and the Flicker of Freedom", in *American Philosophical Quarterly*, 36: 241-251.

Graham, Peter A., 2017, "The Epistemic Condition on Moral Blameworthiness: A Theoretical Epiphenomenon", in Philip Robichaud, Jan Willem Wieland, eds., *Responsibility: The Epistemic Condition*, Oxford University Press, Oxford: 163-179.

Hieronymi, Pamela, 2006, "Controlling Attitudes", in *Pacific Philosophical Quarterly*, 87: 45-74.

Khoury, Andrew C., 2018, "The Objects of Moral Responsibility", in *Philosophical Studies*, 175: 1357-1381.

Mele, Alfred R., 1992, *Springs of Action: Understanding Intentional Behavior*, Oxford University Press, Oxford.

Mele, Alfred R., 2017, *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*, Oxford University Press, Oxford.

Pereboom, Derk, 2001, *Living without Free Will*, Cambridge University Press, Cambridge.

Zimmerman, Michael J., 2002, "Taking Luck Seriously", in *The Journal of Philosophy*, 99: 554-576.

Zimmerman, Michael J., 2006, "Moral Luck: A Partial Map", in *Canadian Journal of Philosophy*, 36: 585-608.

In memoriam
Paolo Parrini

Introduction
Paolo Parrini & relative *a priori* principles

Kenneth R. Westphal

1. *To commence, and to commemorate*

Paolo Parrini devoted concerted philosophical attention to theoretical philosophy, re-examining core issues in epistemology, philosophy of language and history and philosophy of science, not only for their intrinsic philosophical interest, but also for their cultural significance. This pair of papers published here in English he himself affiliated closely. "Analyticity and Epistemological Holism: Prague Alternatives" appeared originally in Italian in 2006; "Quine on Analyticity and Holism. A critical appraisal in dialogue with Sandro Nannini," in 2018. He translated both into English early in 2020, posting their original Italian together with their new English versions on his own website.[1] Doubtless both are related to his research interests in Herbart's conceptual *Bearbeitung*, which surely must be a vigorous form of conceptual explication. Most unfortunately, Paolo was taken from us suddenly, unexpectedly, at the start of July (2020). What more we can learn from him, we shall learn from his considerable published accomplishments. This brief Introduction seeks to epitomize the core issues and significance of this pair of papers, in tribute to him and his very substantial philosophical achievements.[2]

2. *Core issues in semantics, epistemology and history & philosophy of science*

Two central theses of Paolo Parrini's thematically linked papers may be stated briefly: (1) There is an important role for those 'conventions' or stipulations involved in setting basic units and procedures of physical measurement, quite distinct to those 'conventions' or principles set in order to frame some

[1] On Academia.edu: <http://unifi.academia.edu/PaoloParrini>.

[2] For broader consideration of Parrini's philosophical views see Parrini (2017), Stöltzner (1998), Oliva (2015), Lanfredini & Peruzzi (2013), (2016), and Lanfredini (2021). For concise, independent explication of relative *a priori* principles under a different designation, see Toulmin (1949).

philinq IX, 1-2021, pp. 59-78
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.360

kind or domain of scientific inquiry, although both may be regarded as species of *relative a priori* principles. (2) Prospects for such species of relative *a priori* principles emerged already in 1934 at the meeting in Prague on scientific philosophy, attended by both Carnap and Quine.

Stating these two theses directly in this way does not yet begin to address why or how Parrini devotes such extensive, careful re-examinations – philosophical, textual and historical – to explicating, elucidating and justifying these two theses. One reason for his great care is that their justification does not fit neatly into any of the typical options; they concern logically contingent principles which are presupposed by specific kinds of scientific, physical inquiries; they are warranted indirectly yet very powerfully by empirical research together with historical assessment of their advent, implementation and alternatives. Another reason for his great care with these theses is that they have been in the public (philosophical) domain for nearly a century, yet have been widely, persistently neglected due to typical philosophical aspirations, programmes and ways of thinking which have pervasively over-simplified the list of (purportedly) relevant options. Parrini's essays are as much about *how* to philosophize well, as about what best to learn philosophically about these issues and from these texts and debates. All this belongs to Parrini's "open-texture[d] rationality" (2021b: 96, *cf. n.*4), and to his demonstration that rational assessment and justification can indeed thrive within an open-textured, fallibilist and far more hermeneutical approach to issues (primarily) in theoretical philosophy.

3. *Carnap's & Parrini's pragmatics*

3.1. I begin at what may seem an incidental point: Parrini's fourth footnote to "Analyticity and Epistemological Holism: Prague Alternatives" (2021a). There he recalls emphatically that Carnap himself took seriously issues in pragmatics (of language) as well as intensional and intentional phenomena, where 'intensions' are meanings or classifications, and 'intentions' are psychological attitudes or aims. Parrini's footnote is no mere historical aside; here readers should ask themselves, *why* is it significant for Parrini to remark upon Carnap's continued concern with intensional, intentional and pragmatic issues? 'Pragmatics' concerns the actual use of language by actual speakers to express statements (whether assertions, queries or imperatives) in various actual contexts, in contrast to syntax (which concerns grammatical structure or formation) and to 'semantics' as concerned with meaning, classification, descriptive 'content' or intension. I stress this sense of 'semantics', because the term is deeply ambiguous between meaning and reference, *qua* connections (if any) between what is said, and any actual individuals about which anything may be said. Parrini's

emphatic recollection of Carnap's concerns with pragmatics goes to the core of Carnap's and also Parrini's issues with Quine, who spent his career seeking to eschew pragmatics, intensions and intentions so far as possible in favor of his Thesis of Extensionalism. Substantiating his Thesis of Extensionalism requires Quine to appeal to the most minimal behaviourist account of speech and language. Countenancing only (formalized) syntax and semantics requires treating reference as nothing other than successful description, *per* Russell's account of definite descriptions. Quine never noted that his favorite example of a putative definite description, 'the shortest spy', may be either empty or ambiguous; ambiguous if the shortest spies are triplets of the very same (physical) stature and clandestine profession, or empty, if ever we have the great fortune that their entire profession vanishes from the face of the Earth. In principle, descriptive specificity cannot suffice for definite reference, whether singular or plural, to any one, nor to any group, of specific individual(s). This basic point about syntax, semantics and pragmatics has far-reaching implications.

3.2. Parrini (2021a: 81) notes a very important point from Carnap's Logical Syntax of Language (1934/1937), part of which I quote here more fully. Carnap states:

> If a sentence of the material mode of speech is given, or, more generally, a sentence which is not a genuine object-sentence, then the translation into the formal mode of speech need not always be undertaken, but it must always be possible. *Translatability into the formal mode of speech constitutes the touchstone for all philosophical sentences,* or, more generally, for all sentences which do not belong to the language of any one of the empirical sciences. (Carnap 1934/1937: §80)

By "transposed (*verschoben*) mode of speech," Carnap means any mode of speech which cannot be construed directly and literally as pertaining to one or more objects (*e.g.*, metaphors, figurative speech), a feature he considers to pervade natural languages and its material (*inhaltliche*) mode of speech.[3] In the quoted passage, Carnap states expressly, indeed stresses, that *not* all sentences can, nor should, be translatable into the formal mode of speech. In particular no genuine object sentences, and hence no such sentences from the sciences,

[3] „Wollen wir die inhaltliche Redeweise unter einen allgemeinen Begriff bringen, so konnen wir etwa sagen, daß sie eine besondere Art von verschobener Redeweise ist. Dabei wollen wir unter einer verschobenen Redeweise eine solche verstehen, bei der man, um etwas über den Gegenstand *a* auszusagen, etwas Entsprechendes über einen Gegenstand *b* aussagt, der zu *a* in einer bestimmten Beziehung steht (das soll keine genaue Definition sein). Jede Metapher ist z. B. eine verschobene Redeweise; aber auch verschobene Redeweisen anderer Art kommen in der üblichen Sprache häufig vor, weit häufiger, als man zunächst glauben mag" (Carnap 1934: §80).

can or should be so translated, nor translatable. This is a very important observation by Parrini, widely disregarded in discussions of Carnap's views. "Genuine object sentences" are genuinely, directly, literally about *objects*. Hence their use, assertion and assessment *require* pragmatics, because they require actual use by actual scientists to state or query anything about actual objects (whether accurately or inaccurately, justifiedly or not).

However, if philosophers only consider formalized syntax and formalized semantics *qua* classifications or intensions or meanings (not reference and not referents, *i.e.*, not *designatae*), then *per force* they only consider sentences metalinguistically, and disregard all pragmatics, all actual use, all actual reference, all actual referents, and any actual statements. This is what Quine did, and what his followers have done, if perhaps inadvertently. Carnap had expressly and repeatedly cautioned about the differences between his formally regimented syntax and semantics, and their proper *use* in any scientific context, in contrast to natural languages, which are far richer and (quite literally) unruly, including in *The Logical Syntax of Language*. Exactly this contrast between formalized syntax and semantics, and any natural language, Quine (1951: 34; 1953: 36) disregarded when contesting analyticity: He expressly sought to understand 'analyticity' within *natural* languages, pointedly dismissing Carnap's formalized languages and disregarding his cautions about natural language! However, Quine's constant recourse to "semantic ascent" puts language permanently on an arid holiday, because in principle it thus lacks all pragmatics, and hence any *real* use in any *actual* context by any actual person. Already in 1932-33 and repeatedly in later publications Carnap stressed that his formalized syntax and formalized semantics *require* their pragmatic counterpart of actual use by actual scientists in actual contexts of actual scientific inquiries to have any *real use* or content.[4] Without pragmatics, formalized syntax and formalized semantics are referentially, cognitively, scientifically empty forms! *Also sprach Rudolf Carnap.*

3.3. The fundamental importance of pragmatics within Carnap's philosophy of language, and to his formalized syntax and semantics, is reinforced by this feature of conceptual explication. Although Carnap first explicated his method of philosophical explication in 1950, he had been using it since at least the *Aufbau* (1928). Both in "Empiricism, Semantics and Ontology" (1950b, rev. 1956) and in his official explication of 'explication' (1950a: 1-18),

[4] Carnap (1932-33): 178, 179, 182; (1942): §5; (1963b): 923, 925-927. This is Carnap's "descriptive semantics," in contrast to "pure semantics," which is his formalised syntax and (after 1942) formalised semantics. Carnap's characterisation of "descriptive semantics" makes plain that it belongs to Morris's pragmatics of actual language use.

Carnap makes plain that however, and however much, they may innovate using formalized syntax or semantics, any conceptual explication, and likewise any linguistic framework, must and can only be *assessed* and either adopted, adapted, rejected or replaced within *actual* contexts of actual use, including the original contexts of use whence came the relevant *explicandae* (*cf.* Westphal 2015a). This point bears emphasis today, for "conceptual engineering" traces back to Carnap's (1950) views; Carnap (1963a: 66, 1963b: 912) himself speaks of "language engineering." Unfortunately, most of today's interest in "conceptual engineering" follows Quine's arid semantic ascent by disregarding the requirement to assess any bit of conceptual engineering by examining its use and usefulness within some *actual* application to address some actual, first-order problem(s) or issue(s).[5]

3.4. A further important precautionary note against excessive semantic ascent affords a friendly amendment to Parrini's view. Michael Friedman's (1983: xv) examination of the foundations of space-time theories may have aimed to support realism, Parrini (2010: 210) noted, yet closer analysis reveals it does not. Friedman's several "Newtonian" models preserve no more than Newton-inspired kinematics, but cannot preserve Newton's dynamics, i.e., his causal theory designed to explain robust, established kinematic regularities by multiple, independent, precise measures of the exact rate of gravitational attractions (pair-wise) across our solar system, including a wide range of terrestrial kinematic phenomena (Harper 2011, cf. Huggett & al 2013). All of Newton's measures require appeal to material mass and its proportional gravitational power attraction; whereas no mention of, nor reference to, 'mass' is preserved by Friedman's elaborate modeling – none at all! Hence it cannot have modeled Newton's dynamics, hence also not Newton's mechanics – at all.

Once I had occasion to remind Friedman of Carnap's insistence that his formalized syntax and formalized semantics are, expressly, abstractions from pragmatics, from actual linguistic usage by actual people to talk about their surroundings, and that without pragmatics, Carnap's formalized syntax and formalized semantics are mere empty forms, as he acknowledged expressly in reply to Zilsel and Duncker (Carnap 1932-33). Friedman blithely dismissed Carnap's pragmatics and the point of my recalling it to his attention. This occasion was subsequent to Friedman (1992), but his explicitly stated preference for Carnap's (purported) "formalism" matches exactly the formalist mistake in

[5] For a representative discussion of "conceptual engineering" today, see Eklund (2015). Rescher (2017) is better about applied use, but neglects Parrini's key point regarding distinct kinds of relative *a priori* principles.

his (1983) modeling of Newton's (purported) theory (*cf.* Westphal 2020: §72). The vagaries of his (2001) purported "dynamics of reason" (*cf.* Parrini 2021b: *n.*11) await in the wings of his meta-linguistic formalism.

4. *Pragmatics & coördination principles*

These points about pragmatics help elucidate Parrini's re-examination of Reichenbach's (1920) "coördination principles" (*Zuordnungsprinzipien*), by identifying what might be called four grades of coördinative involvement.

4.1. A first grade of coördinative involvement concerns anyone's coördinating any one thought with any one actual individual by *ascribing* what one thinks *to* that individual, which one indicates deictically as being right there and then.[6] Misdescription or mistaken attribution are compatible with successful deictic reference to an extant, intended, indicated individual (*per* Donnellan 1966, Evans 1975). This may not sound exciting, but it underscores the crucial importance of pragmatics of language; this deictic point is also central to Wittgenstein's point that no map can indicate its own scope of reference (where its own 'territory' lies). Even a diagrammatic map tucked into a corner showing the area of the main map can be used only if one can identify which area *of the planet* is that area within which the diagram marks out the specific area represented by the main map.[7]

4.2. A second grade of coördinative involvement concerns indicating where & when or there & then – *i.e.* individuating and identifying specific occasions in specific locations – by using a coördinate system of spatial and temporal relations to designate any relevant occasion(s) or location(s) of any designated, indicated individual(s). The issues involved in understanding and using such coördinate systems, however informal or commonsensical they may be, are complex. The key point here is that in principle they cannot be addressed by empiricist accounts of conceptual content (intension). Neither Hume nor Carnap (1928) can account for the intension (meaning) of our commonsense conceptions of 'time', 'times' or 'occasions', nor for our capacity to use these conceptions to identify anything as occurring before, during or after anything else (Westphal 1989: 230-232 (*n.*99); 2103). Regarding our conceptions of 'space', 'spaces' and

[6] Outside philosophy, the referential phenomena philosophers consider under the headings of 'indexical' or 'demonstrative' expressions or gestures are known by the Attic Greek term *deixis* (Bohnemeyer 2015).

[7] This first level of coördinative involvement for deictic reference to particulars is central to what I call "cognitive semantics," which Parrini (2021b: *n.*9) generously endorses.

'spatial location(s)', Howard Stein (1967) noted that Newton recognized that Descartes' official conception of space in terms of nothing but contiguity relations cannot possibly account for our capacity to identify even the simplest kinds of trajectory, because once any one particular departs from the vicinity of its immediately contiguous neighbors, the spatial location it had occupied *literally* no longer exists![8] Tracking trajectories, including our own local motions, is required (presupposed) by the first grade of coördinative involvement (§4.1). Kant is thus correct that our *concepts* of 'space', 'spaces', 'time' and 'times' must be *a priori*; they cannot be defined, acquired or learned by our sensing particulars around us. However, his transcendental idealist "explanation" of how we can have those *a priori* concepts fails (Westphal 2004: §§27-28). This is the point underlying Parrini's frequent dismissal of *a priori* "forms of intuition" (*cf.* Parrini 2021b: 81, whilst nevertheless advocating relative *a priori* framework principles – and hence, whatever concepts or conceptions are required to frame, formulate, understand, use or assess such frameworks.

4.3. A third grade of coördinative involvement concerns specifying measurement procedures and metrics within engineering and the exact sciences. Too often it is supposed that metrics can be set merely by convention, or perhaps by convention plus technique or technical apparatus (*i.e.*, observational, measurement or experimental devices). This is too glib. To be at all useful, measurements must be regular, reliable and informative. Neither independently nor conjointly do convention, theory or device suffice to establish measurement metrics. This is because no measurement procedure (including its affiliated conventions, theory or devices) can establish whether any unknown natural phenomenon happens, unbeknownst to these *calculatores*, to interfere with their use, results or interpretation of that procedure. This is a crucial reason why measurement procedures must be understood as involving relatively *a priori* – logically contingent, hence fallible, revisable and with care also corrigible – principles, which cannot be defined or justified merely empirically.[9]

4.4. A fourth grade of coördinative involvement concerns an especially important case of the third: Reichenbach's *Zuordnungsprinzipien* in his original

[8] See Westphal (forthcoming): §4. Stein is amongst Parrini's (2010: 192, 103) sources, too.

[9] This point was first brought to my attention by Laymon (1991: 173-177), with whom I had studied Carnap's semantics. Yet I did not then know enough to appreciate properly Laymon's paper. Parrini's re-examination of these issues led me back to Laymon, and prompted my renewed appreciation of both their views (Westphal 2015b: §3.2). I'm grateful to Paolo for his important reminder, and also for his gracious endorsement of my re-presentation of his account of this important point (Parrini 2021b: *n.*11).

account of relativity theory. In 1920 Reichenbach rightly recognized that the specifications of simultaneity within relativity theory cannot be merely conventional; that (for reasons indicated in §4.3) the specifications of, and metrics for, 'simultaneity' require a physical postulate of a natural regularity, which is presupposed, but cannot be directly evidenced, by relativity theory nor its related observations and their theoretical analysis. Unfortunately, the conventionalist orthodoxy of his logical empiricist colleagues led him to rescind his correct understanding of these fundamental *Zuordnungsprinzipien* within relativity theory to rejoin their conventionalist fold. This is exactly the point central to Parrini's (2010) distinguishing – repeatedly, pointedly and correctly – between the 'conventionalism' (truly a misnomer) involved in astrochonometrical measurement procedures, and whatever conventions may be stipulated at a higher level of theoretical generality (e.g., Carnap's *L*- and *P*-rules). Both levels involve relatively *a priori* concepts and principles, though of importantly different kinds: framework principles and measurement metrics.

All four grades of coördinative involvement require *pragmatics*: actual linguistic use by actual persons (including scientists) in actual circumstances who actually perceive, investigate or measure some of their actual surroundings (including their equipment). None of this can come properly into view from Quine's lofty logical point of view; neither can it come into view from any merely formalist modeling – as evident in re-examining Friedman's (1983) purported foundations of space-time theories. Why call these four grades of coördinative *involvement*? Because each is an important involvement *with the world*, with actual contexts, actual phenomena, actual problems and actual co-inquirers. Mere logical possibilities need not apply; mere intensions, no matter how good, do not suffice for any real use to address any real issue, not in philosophy, nor in allied disciplines.

5. *Explanatory desiderata, Newton's rules of method & testing coördination principles*

Newton's methods and procedures exhibit the kind of relative *a priori* principles central to Parrini's empirical realism, and indeed Newton's methodological Rule 4 provides a very powerful test for scientific explanations using distance forces. In contrast, the control or testing of measurement procedures and metrics, especially those used at micro or macro (astronomical, cosmological) scales, are ill-understood by appeal to the "theoretical desiderata" or "explanatory virtues" touted by empiricists: simplicity, unity, comprehensiveness, precision, elegance, unifying power or fecundity. Empiricism with its entirely descriptive aspirations, its regularity account of causality and its covering

law account of explanation is suited only to descriptive kinematics, not to any causal dynamics which can *explain* kinematic regularities.

Despite some nomenclature in the *Principia*, Newton's mechanics does not have the quasi-axiomatic structure central to empiricist philosophy of science, according to which high-level theoretical statements (instances of Carnap's *L*- and *P*-rules) are linked to observational or experimental reports by intermediate-level correspondence rules. Instead, Newton's Books I and II develop a precise and powerful mathematical approach to *measuring* gravitational attractions (in Book III) pair-wise across our solar system and on Earth. Newton's Definitions 5-7 define *quantities*, expressly they define *measures*, of specific kinds of forces; they do not define *forces*. Newton's methods provide independent, precise, robust methods of successive approximation through elimination of initial idealizations by reiterated use of the very same explanatory resources, in order to use carefully measured, observed *kinematic* phenomena to *measure* very precisely *forces* of gravitational attraction. Empiricist methods, including Glymour's "boot strap" method, cannot suffice to achieve what Newton's methods *did* achieve: to disentangle weight from mass of orbiting bodies. These and many more important results have been achieved by work spear-headed by Howard Stein, culminating in Harper (2011; *cf.* Harper 2020).

Newton's Rule Four of experimental philosophy states:

> In experimental philosophy, propositions gathered from phenomena by induction should be considered either exactly or very nearly true notwithstanding any contrary hypotheses, until yet other phenomena make such propositions either more exact or liable to exceptions. (Newton 1999: 796; 1726: 389)

Newton directly adds, "This rule should be followed so that arguments based on induction may not be nullified by hypotheses" (*ibid.*). Newton's Rule Four requires any competing scientific hypothesis to have, not merely empirical evidence in its favor, but sufficient and sufficiently *precise* evidence to make an accepted scientific hypothesis either "more exact" or to restrict it by demonstrating actual "exceptions" to it.

As scientific inquiries are extended to ever greater – or conversely, to increasingly microscopic – scales, the measurement procedures and metrics used in established sciences are tested ever more severely, providing ample opportunities to corroborate them through continuing adequate performance at ever greater extremes of precision, or to adapt or replace them to improve upon their detected, no longer sufficient precision. This in fact was achieved by Einstein's theories of relativity, *by using Newton's methodology!* Fed Einstein's new, more precise data and analysis, Newton's methodology strongly *favors* relativity theory over classical mechanics (Harper 2011: 378-385, 392, 394-396).

According to Steen Brock (2003), this same methodology pertains also to the successful development of quantum mechanics within the Helmholzian methodological tradition.

An important feature of Newton's Rule Four is that it too is rooted in the deictic point that to be at all relevant scientifically, a competing hypothesis must *have* empirical evidence which supports it, and indeed, supports it *differentially* in contrast to an established theory. This requires that the competing hypothesis *is* referred to actual, identified (and relevant) natural phenomena; 'referrability in principle' – mere intension – does not suffice to *have* a truth-value, nor value as an approximation – nor even to be erroneous! This deictic requirement of Newton's Rule Four is directly and powerfully supported by Kant's semantics of singular cognitive reference. These scientific and methodological findings are among Parrini's (2021b: *n.*10) reasons for having so endorsed my account of Kant's semantics of singular cognitive reference. Kant's semantics of singular cognitive reference demonstrates that mere logical possibilities have *no* cognitive status whatever unless and until they are *referred* in specific ways by specific people to identified, localized relevant individuals (at which point they are no longer mere logical possibilities). This is the important point required by the first grade of coördinative involvement discussed above (§4.1), which requires the second grade as well (§4.2). The developments in physics, especially astronomy, though also high-energy physics and quantum mechanics, require the third and fourth grades.

These results – which can only be stated briefly here[10] – show that, and when detailed, show *how*, robust testing of relatively *a priori* framework and also measurement principles and practices is possible, consistently with justificatory fallibilism, with various kinds of semantic holism *and* with Duhem's epistemological point about in ineluctable use of a host of theoretical as well as experimental or observational resources when testing any one scientific hypothesis. This is the central point of Parrini's demonstration that rational assessment and justification can indeed thrive within an open-textured (2021b: 96, *cf. n.*4), fallibilist and far more hermeneutical approach to issues of assessment and justification within theoretical philosophy, by distinguishing and carefully integrating the host of relevant factors, rather than conflating them with one or another theoretical over-simplification. This marks a key tension between philosophical aspirations to manage everything merely formally, despite the insufficiency *in principle* of formal techniques for sub-

[10] For details, please see Westphal (2014) or (2020): §§66-74 (rather than the paper Parrini (2021b: *n.* 12) cites as 'forthcoming', which remains under review).

stantive philosophy,[11] and aspirations to adequacy and understanding, which require hermeneutical care and self-criticism. *This* is the key "parting of the ways" within contemporary philosophy.

Parrini adroitly observes that the havoc wrought by Kuhnian paradigm shifts upon the logical empiricist account of the (tri-level) structure of scientific theories

> …added extra value to the thesis – characteristic also of Carnap's epistemology with its distinction between internal and external questions – according to which we cannot understand the structure of and changes within science without taking into consideration the presuppositions that provide the framework for scientific activity. (Parrini 2021a: 87)

Parrini rightly insists that these presuppositions are of various distinct kinds, a point obscured by Kuhn's "paradigms," and not much improved by his later terms, "exemplar" and "disciplinary matrix" – largely because Kuhn treated these as integrated packages of methodology + theory + findings, where changing any one of these components required systematically changing the others. In this regard, Kuhn's holism was cut of the same entirely intensional, meta-linguistic, merely descriptive cloth as Quine's.[12]

6. *Quine on analyticity & holism. A critical appraisal*

In his first introductory section, Parrini (2021b) explains very well his aims regarding Professor Nannini's views. The body of Parrini's paper (§§2-4) is a concise *tour de force* review and critical assessment of Quine's chronic prevarications regarding holism and the distinctions between analytic/synthetic and between *a priori/a posteriori*. Quine's pervasive prevarications all result from overly simplistic theses and options, none of which is immune to criticism or replacement. Quine's views on these fundamental points deserve such re-examination and re-assessment because they remain pervasive and apparently persuasive, official disavowals by many more recent philosophers not

[11] Formal techniques can be used to specify (Carnapian) "meaning postulates" or other relatively *a priori* concepts or principles, but formal techniques alone do not suffice to *assess* those postulates or principles, nor their proper, effective *use* in connection with any *actual* problem or its actual context. (This is not to reject formal techniques; it only cautions about their proper understanding and effective use.)

[12] In these regards, Rorty's neo-pragmatism and Putnam's internal realism are entirely within Quine's meta-linguistic, merely intensional ambit. The "strong programme" in sociology of knowledge joins their ranks by pointedly disregarding issues of truth, evidence or justification, because those issues are scientific, and so (admittedly) lie beyond the competence of sociologists of science.

withstanding. More noteworthy, both philosophically and historically, is that Carnap had already developed more sophisticated and cogent alternatives to Quine's views, which were neglected due to the popular success of Quine's publication campaign. Quine always sought simplicity, yet persistently disregarded Einstein's (2000: 314) precisification of Okham's razor: Everything must be made as simple as possible, but no simpler. – Accuracy, adequacy and insight are far more important. The fundamental issues involved regarding language, knowledge and science require at least the care Carnap devoted to them, yet Parrini – also in this pair of papers – augments Carnap's views cogently. One important point Parrini rightly stresses is that whatever *semantic* holism (regarding meaning or intension) may characterize language, or languages, or large domains of specific languages, such holism is distinct to the kind of much more limited *cognitive* "holism" involved in testing scientific principles, hypotheses or explanations characterized by Duhem. Duhem's thesis is specifically cognitive or epistemological, not semantic (in the sense of meaning or intension) (Parrini 2021b: §4). The fallibility of fundamental principles and procedures, due to their logical contingency, can only be linked to their credible corrigibility by careful attention to the complexity and complex interrelations of semantic, epistemic, measurement and experimental or observational methods, techniques and findings. These links are in part historical, insofar as our current best options must be known to be superior to their available alternatives, both historical and contemporaneous (*per* Rule Four); in part these links are current (contemporaneous), insofar as they structure and guide effective inquiry, analysis and findings; and in part these links are anticipatory: they are and remain justified unless and until an actual alternative succeeds in providing robust results with improved accuracy. Only in this way is fallibilism consistent with cognition, with our actually knowing – if provisionally or approximately – whatever we do about nature or, *mutatis mutandis*, within other domains of inquiry. The relevant relative *a priori* principles cannot be merely linguistic, nor merely meta-linguistic; they are substantive principles subject to empirical assessment – if only through long-term, large-scale scientific explorations and successes (Parrini 2021b: §5).[13]

Parrini raises these issues about Quine to Professor Nannini in connection with contemporary naturalism, especially Quine's naturalism. To put the point

[13] To lump these results, Newton's methodology or Newton's use of Rule Four under the heading of theoretical 'fecundity' would be to cover everything by saying nearly nothing; it would utterly evacuate the content of Newton's philosophical and scientific achievements merely to preserve empiricist's meta-linguistic preferences. What appear to lie beyond empirical control are empiricists's chronic misconceptions and over-simplifications of empirical science. Robust history and philosophy of science (HPS) can do and has done much, much better.

as briefly as possible: None of the phenomena illuminated by Parrini's explication of various kinds of relative *a priori* principles and their humanly possible use, summarized in the preceding sections (§§2-6), can be understood merely naturalistically, and especially not by Quine's naturalism. One key reason for this shortcoming Quine himself highlighted in 1936, in "Truth by Convention," in which he demonstrated (*inter alia*) that no formally stated rules can be used to first specify any marks as signs, nor any series of marks as statements or as rules, without someone's properly *using* by thinking and acting (writing) in accord with, and on the basis of, *intensions* which structure her or his *intentions* so to think and write, so as to *state* (not merely scrawl or mark) basic symbols, formation rules or transformation rules for any formally specified logistic system, however elementary. Quine never recognized that this fundamental point refutes in advance his cherished Thesis of Extensionalism![14] Conversely, none of these syntactic, semantic and pragmatic points against naturalism provide the least aid or support to rationalism, nor to idealism and certainly not to anti-realism nor to scepticism! Parrini's view is indeed a robust empirical realism, undogmatic, subtle and supple.[15]

The broad yet also deep and incisive account of these substantive versions of relatively *a priori* principles and their roles in scientific knowledge Parrini developed is exemplary in yet another important regard. Issues in epistemology, semantics or philosophy of science are often regarded as technical matters for various specialists; non-specialists often make do with simplified accounts of these issues or views. Such a view is often taken of issues or views associated with logical positivism or logical empiricism. Not only Parrini's own research, but his extensive and intensive involvements with both the Pittsburgh HPS programed and also the Western Ontario group (whose series includes Parrini 1998), belies such convenient caricatures. Both groups demonstrated serious, sustained interest in Parrini's research, extending over decades. Their interest corroborates the international calibre of Parrini's philosophical research.

The willingness to settle for simplifications rather than accuracy has become a prevalent philosophical vice (*cf.* Parrini 2019; Addis & Westphal 2019), one fostered by the kinds of over-simplifications characteristic of Quine's publications, some of which are diagnosed in this pair of Parrini's

[14] For detailed explication and defence of this strong claim, see Westphal (2015a); I am very gratified that Paolo (2021b: *n.*5) concurs with the substance of my analysis, and am happy it brings me into accord with his own earlier work. (I had read his (1976), with keen interest, benefit and pleasure – he kindly gave me a copy of it and of his (1983) some years ago – but I began my (2015a) from scratch, to be as thorough and as strictly internal as possible when presenting and assessing Quine's semantics.)

[15] The soundness and significance of Parrini's highlighting Carnap's pragmatics in connection with scientific realism is corroborated by Salmon's (1994) comment on Parrini (1994).

papers, which demonstrate, as a matter of public record, that significantly more adequate views have been available all along, not least from Carnap. Such is the price of promoting simplicity over adequacy. Carnap was indeed the most sophisticated logical empiricist. By so carefully re-thinking these issues, texts and debates, Paolo Parrini has philosophized with Carnap in order to improve still further by identifying key roles of relative *a priori* principles within the physical sciences.

7. *Characterizing this occasion*

Finalizing Paolo's papers for publication here, *in memoriam*, has come to me by this route. I first learned of Paolo and his interests in Kant's epistemology when I came upon his collection, *Kant and Contemporary Epistemology* (1994). I wrote him about it; he kindly replied, indicating his research interests had shifted focus. Subsequently I found his *Knowledge and Reality* (1998), and noticed that his shift in research interests had not left Kant behind, but rather he too sought to develop a credible and illuminating account of a substantive relative *a priori*, which is neither merely linguistic nor merely meta-linguistic. Our correspondence developed substantially by 2012, when Paolo kindly took interest in some of my recently published research on our shared interests. We met personally in 2015, when Cinzia Ferrini (Trieste) generously arranged a seminar presentation of a volume I edited, *Realism, Science and Pragmatism* (Routledge 2014), on which Paolo generously commented. Our exchange and our replies to questions from the seminar were kindly edited and published by Ferrini (2015) in a special issue of *Esercizi Filosofici*. Our correspondence intensified further, as we discovered our considerable convergences, not merely on topics and resources, but in substantial philosophical findings.

When Paolo posted this pair of papers on his website I happened to be on line, and immediately saw notice of his new post. I downloaded the papers promptly, read them both, and offered to polish his prose—*no* arduous undertaking, to be sure! Paolo and I share a keen interest in polished philosophical prose, which requires native competence; Paolo's philosophical views and prose deserve no less. I had already made much of the (minor, entirely stylistic) revisions when suddenly these papers became suited to this memorial commemoration. It is personally a great honor to me to provide him this one further kindness. His philosophical care and insight are matched by his unfailing modesty and engagement with substantive philosophical issues and interlocutors. Philosophically, I dearly wish we could have learnt what he sought to make of Herbart's conceptual *Bearbeitung*. Nevertheless, he and

his philosophy have enriched my own. His sudden, unexpected loss remains a great burden to all who knew him, and to me personally. Thank you, Paolo, for these and for so many more reasons, some of which I shall yet discover in your rich philosophical research![16]

Kennet R. Westphal
westphal.k.r@gmail.com
Boğaziçi Üniversitesi, Istanbul

References

Paolo Parrini

1976, *Linguaggio e teoria. Due saggi di analisi filosofica*, La Nuova Italia, Firenze.

1983, *Empirismo logico e convenzionalismo*, Franco Angeli, Milano.

1994, "With Carnap Beyond Carnap. Metaphysics, Science, and the Realism/Instrumentalism Controversy", in Wesley Salmon and Gereon Wolters, eds., *Logic, Language, and the Structure of Scientific Theories*, *Proceedings of the Carnap-Reichenbach Centennial, University of Konstanz, 21–24 May 1991*, University of Pittsburgh Press/Universitätsverlag Konstanz, Pittsburgh/Konstanz: 255-277.

1994, ed., *Kant and Contemporary Epistemology*, Kluwer, Dordrecht.

1995, *Conoscenza e realtà. Saggio di filosofia positiva*, Laterza, Roma-Bari.

1998, *Knowledge and Reality. An Essay in Positive Philosophy*, Kluwer, Dordrecht; Eng. trans. by Paolo Baracchi with revisions and additions to *idem*, 1995.

2006, "Analiticità e olismo epistemologico: alternative praghesi", in Rosa Calcaterra, ed., *Le ragioni del conoscere e dell'agire. Scritti in onore di Rosaria Egidi*, Franco Angeli, Milano: 190-204.

2007, "Il convenzionalismo epistemologico al di là dei problemi geocronometrici", in *Rivista di filosofia*, 62, 4: 711-741.

2010. "Epistemological Conventionalism beyond the Geochronometrical Problems", in Mario De Caro and Rosaria Egidi, eds., *The Architecture of Knowledge: Epistemology, Agency, and Science*, Carocci, Roma: 191-223 (Orig. It. version: *idem*, 2007; It. rev. version: *idem*, 2011).

[16] I am grateful to the editors of *Philosophical Inquiries* for accepting my suggestion that these two papers be published here in tribute to Paolo Parrini, and for their generous invitation to provide this substantive Introduction. To Simonetta and their children I reiterate my deepest condolences; I thank Simonetta for her very kind and generous correspondence. Again I thank Cinzia Ferrini for so kindly checking my Introduction and for intermediating my correspondence with Simonetta and also with Roberta Lanfredini, whom I thank for very kindly checking this Introduction in advance. (This research and editorial work has been supported in part by the Boğaziçi Üniversitesi Research Fund; BAP; grant code: 18B02P3.)

2011, *Il valore della verità*, Guerini e Associati, Milano.

2015, "Empirical Realism without Transcendental Idealism. Comment on Kenneth R. Westphal", in Cinzia Ferrini, ed., *Approaching Contemporary Philosophical Problems Historically: On Idealisms, Realisms, and Pragmatisms*; proceedings of the Workshop, "Kenneth R. Westphal & Paolo Parrini: A Real Dialogue on an Ideal Topic", in *Esercizi Filosofici*, 10, 1: 1-96, 41-61, 81-96.

2017, "Science, Knowledge, Rationality, and Empirical Realism. A Conversation with Paolo Parrini", in Silvia Benso, ed., *Viva Voce. Conversations with Italian Philosophers*, SUNY Press, Albany NY: 185-199.

2018, "Quine su analiticità e olismo. Una valutazione critica in dialogo con Sandro Nannini," in Christoph Lumer and Giacomo Romano, eds., *Dalla filosofia dell'azione alla filosofia della mente. Riflessioni in onore di Sandro Nannini*, Corisco Edizioni, Roma-Messina: 95-116.

2019, "Philosophy Today: Cries of Alarm and Prospects of Progress", in *SATS—Northern European Journal of Philosophy*, 20, 2: 1-20.

2021a, "Analyticity and Epistemological Holism: Prague Alternatives", in *Philosophical Inquiries*, 9, 1: 79-93.

2021b, "Quine on Analyticity and Holism. A critical appraisal in dialogue with Sandro Nannini", in *Philosophical Inquiries*, 9, 1: 95-112.

Other authors

Bohnemeyer, Jürgen, 2015, "Deixis", in J.D. Wright, ed., *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed., Elsevier, Amsterdam, 6: 52-57.

Brock, Steen, 2003, *Niels Bohr's Philosophy of Quantum Physics in the Light of the Helmholtzian Tradition of Theoretical Physics*, λογος, Berlin.

Carnap, Rudolf, 1928, *Der logische Aufbau der Welt*, Weltkreis, Berlin, 2nd ed.: 1961; Eng. trans.: *idem*, 1969.

Carnap, Rudolf, 1932-33, „Erwiderung auf die vorstehenden Aufsätze von E. Zilsel und K. Duncker", in *Erkenntnis*, 3: 177-188.

Carnap, Rudolf, 1934, *Logische Syntax der Sprache*, Springer, Vienna.

Carnap, Rudolf, 1937, *The Logical Syntax of Language*, Smeaton, Anne, tr., Kegan Paul, Trench & Truebner, London; rpt.: Littlefield, Adams & Co., Paterson NJ., 1959; Eng. trans. of *idem*, 1934.

Carnap, Rudolf, 1942, *Introduction to Semantics*, Harvard University Press, Cambridge MA.

Carnap, Rudolf, 1947a, *Meaning and Necessity*, University of Chicago Press, Chicago, 2nd rev. ed.: 1956a.

Carnap, Rudolf, 1950a, *Logical Foundations of Probability*, University of Chicago Press, Chicago.

Carnap, Rudolf, 1950b, "Empiricism, Semantics, and Ontology", in *Revue International de Philosophie*, 4: 20-40; 2nd rev. ed. in *idem*, 1956a: 205-221.

Carnap, Rudolf, 1952, "Meaning Postulates", in *Philosophical Studies*, 3, 5: 65-73; rpt. in *idem* 1956a: 222-229.

Carnap, Rudolf, 1955, "Meaning and Synonymy in Natural Languages", in *Philosophical Studies*, 6, 3: 33-47; rpt. in *idem*, 1956a: 233-247.

Carnap, Rudolf, 1956, *Meaning and Necessity*, 2nd rev. ed., University of Chicago Press, Chicago.

Carnap, Rudolf, 1963a, "Intellectual Autobiography", in Paul Schilpp, ed., *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, LaSalle ILL: 3-84.

Carnap, Rudolf, 1963b, "Replies and Systematic Expositions", in Paul Schilpp, ed., *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, LaSalle ILL: 859-1013.

Carnap, Rudolf, 1969, *The Logical Structure of the World*, George, Rolf, tr., University of California Press, Berkeley.

Donnellan, Keith, 1966, "Reference and Definite Descriptions", in *The Philosophical Review*, 75, 3: 281-304.

Einstein, Albert, 2000, *The Expanded Quotable Einstein*, Alice Calaprice, ed., Princeton University Press, Princeton.

Eklund, Matti, 2015, "Intuitions, Conceptual Engineering and Conceptual Fixed Points", in Chris Daly, ed., *The Palgrave Handbook of Philosophical Methods*, Palgrave/Macmillan, London: 363-385.

Evans, Gareth, 1975, "Identity and Predication", in *Journal of Philosophy*, 72, 13: 343-363; rpt. in: *idem*, 1985: 25-48.

Evans, Gareth, 1985, *Collected Papers*, The Clarendon Press, Oxford.

Ferrini, Cinzia, ed., 2015, *Approaching Contemporary Philosophical Problems Historically: On Idealisms, Realisms, and Pragmatisms*; proceedings of the Workshop, "Kenneth R. Westphal & Paolo Parrini: A Real Dialogue on an Ideal Topic", in *Esercizi filosofici*, 10, 1: 1-96, <http://www.openstarts.units.it/dspace/handle/10077/11908>.

Friedman, Michael, 1983, *Foundations of Space-Time Theories*, Princeton University Press, Princeton NJ.

Friedman, Michael, 1992, *Kant and the Exact Sciences*, Harvard University Press, Cambridge MA.

Friedman, Michael, 2001, *Dynamics of Reason*, CSLI Publications, Stanford CAL.

Harper, William, 2011, *Isaac Newton's Scientific Method: Turning Data into Evidence about Gravity and Cosmology*, Oxford University Press, New York.

Harper, William, 2020, 'Newton's Scientific Method', in Eric Schliesser and Chris Smeenck, eds., *The Oxford Handbook of Newton*, Oxford University Press, Oxford, DOI: 10.1093/oxfordhb/9780199930418.013.11.

Hugget, Nick, George E. Smith, David Marshall Miller and William Harper, 2013, "On Newton's Method", in *Metascience*, 22, 2: 215-246; DOI: 10.1007/s11016-013-9745-y.

Lanfredini, Roberta, and Peruzzi, Alberto, eds., 2013, 2016, *A Plea for Balance in Philosophy: Essays in Honour of Paolo Parrini*, 2 vols., Edizioni ETS, Pisa.

Lanfredini, Roberta, 2020, "Il valore della filosofia. La *terza via* di Paolo Parrini", in *Iride. Filosofia e discussione pubblica*, xxxiii, 91: 1-13.

Laymon, Ronald, 1991, "Thought Experiments by Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and as Semantic Domains", in Tamara Horowitz, and Gerald Massey, eds., *Thought Experiments in Science and Philosophy*, Rowman & Littlefield, Lanham, MD: 167-191.

Oliva, Luca, 2015, "Review Esssay: Paolo Parrini's Il valore della verità", in *Kantian Reivew*, 20, 3: 467-478.

Newton, Isaac, 1726, *Philosophiae naturalis principia mathematica*, 3. rev. ed., London; rpt. Glasgow, Maclehose, 1871.

Newton, Isaac, 1999, *The Principia: Mathematical Principles of Natural Philosophy*, Cohen, I. Bernard, & Whitman, Anne, trs., with assistance from Budenz, Julia, University of California Press, Berkeley CAL.

Quine, W.V.O., 1951, "Two Dogmas of Empiricism", in *Philosophical Reivew*, 60, 1: 20-43; rpt. in *idem*, 1953: 20-46.

Quine, W.V.O., 1953, *From a Logical Point of View*, Harvard University Press, Cambridge MA; rpt. Harper & Row, New York, 1961.

Reichenbach, Hans, 1920, *Relativitätstheorie und Erkenntnis apriori*, Julius Springer, Berlin.

Rescher, Nicholas, 2017, "Philosophy as Rational Systematization", in D'Oro, Giuseppina, & Overgaard, Søren, *The Cambridge Companion to Philosophical Methodology*, Cambridge University Press, Cambridge: 32-43.

Salmon, Wesley, 1994, "Comment: Carnap on Realism", in Wesley Salmon, and Gereon Wolters, eds., *Logic, Language, and the Structure of Scientific Theories*, *Proceedings of the Carnap-Reichenbach Centennial, University of Konstanz, 21–24 May 1991*; University of Pittsburgh Press and Universitätsverlag Konstanz, Pittsburgh and Konstanz: 279-286.

Stein, Howard, 1967, "Newtonian Space-Time", in *The Texas Quarterly*, 10, 3: 174-200; rpt. in Robert Palter, ed., *The Annus Mirabilis of Sir lsaac Newton, 1666–1966*, MIT Press, Cambridge MA, 1970: 258-284.

Stöltzner, Michael, 1998, Review of P. Parrini, *Conoscenza e realtà*, in Werner Leinfellner and Eckehart Köhler, eds., *Game Theory, Experience, Rationality: Foundations of Social Sciences, Economics and Ethics in honor of John C. Harsanyi*, Springer Science & Business Media, n.p.: 415-420.

Toulmin, Stephen, 1949, "A Defence of Synthetic Necessary Truth", in *Mind*, 58, 230: 164-177.

Westphal, Kenneth R., 1989, *Hegel's Epistemological Realism*, Kluwer, Dordrecht.

Westphal, Kenneth R., 2004, *Kant's Transcendental Proof of Realism*, Cambridge University Press, Cambridge.

Westphal, Kenneth R., 2013, "Hume, Empiricism and the Generality of Thought", in *Dialogue: Canadian Journal of Philosophy/Revue canadienne de philosophie*, 52, 2: 233-270.

Westphal, Kenneth R., 2014, "Hegel's Semantics of Singular Cognitive Reference, Newton's Methodological Rule Four and Scientific Realism Today", in *Philosophical Inquiries*, 2, 1: 9-65.

Westphal, Kenneth R., 2015a, "Conventionalism and the Impoverishment of the Space of Reasons: Carnap, Quine and Sellars", in *Journal for the History of Analytic Philosophy*, 3, 8: 1-66.

Westphal, Kenneth R., 2015b, "Some Replies to Professor Parrini, to Students and to Members of the Audience", in *Esercizi Filosofici*, 10, 1: 63-79; <https://www.openstarts.units.it/dspace/handle/10077/11912>.

Westphal, Kenneth R., 2018, *Grounds of Pragmatic Realism: Hegel's Internal Critique & Transformation of Kant's Critical Philosophy*, Brill, Leiden & Boston, USA.

Westphal, Kenneth R., 2020, *Kant's Critical Epistemology: Why Epistemology must Consider Judgment First*, Routledge, Oxon & New York.

Westphal, Kenneth R., (forthcoming), *Kant's Transcendental Deduction of the Categories: Critical Re-Examination, Elucidation & Corroboration. Kant's Revised Second (b) Edition (1787), German Text with New Parallel Translation, for Students, Cognitive Scientists, Philosophers & Specialists.*

Wick, Warner, 1951, "The 'Political' Philosophy of Logical Empiricism", in *Philosophical Studies*, 2, 4: 49-57.

Analyticity and Epistemological Holism: Prague Alternatives[1]

Paolo Parrini[†]

Abstract: In the early 1930's Carnap and Quine met in Prague and discussed logic and philosophy. Carnap was working on the *Logische Syntax der Sprache*; when Quine went back to Harvard he published "Truth by Convention." The purpose of the present paper is to establish three main points: (1) in "Truth by Convention" some important aspects of the future position Quine will assume about the analytic/synthetic and the *a priori/a posteriori* dichotomies are already expressed; (2) in the *Logische Syntax der Sprache*, Carnap maintains the distinction between *L*-rules and *P*-rules, at the same time being aware of the holistic character of empirical control and of the possibility to revise the acceptance of every kind of sentences; (3) Quine's idea that the holistic conception requires completely abandoning the analytic/synthetic and *a priori/a posteriori* distinctions does not seem wholly correct. On the contrary, in the *Logische Syntax* Carnap takes a step forward in his conception of the "relativized *a priori*". Thus, we can say that in the Prague years two alternative accounts of the theory/experience relation began to emerge. These two alternatives are still pivotal in contemporary epistemological debate.

In 1934 Carnap and Quine met in Prague and held lengthy discussions on philosophy. Carnap had been in Prague since 1931. From 1926 until the Summer of 1931 he had lived in Vienna working as "instructor of philosophy" at the University with Schlick and taking active part in the Wiener Kreis meetings. In Vienna, Carnap had also met the physicist Philipp Frank, who, at the time, taught at the German University in Prague, succeeding Einstein. It was Frank who helped Carnap obtain the chair of Natural Philosophy that Frank himself had managed to create in his University (see Carnap 1963a: 3, 32).

Quine, instead, arrived in Prague at the end of 1933 to spend the first semester of 1934 there thanks to Harvard's Sheldon Travelling Fellowship which

[1] Essay presented at the international conference, *Philipp Frank: Wien-Prag-Boston* (Vienna, 27–28 Sept. and Prague 30 Sept., 1 Oct. 2004). Italian text, "Analiticità e olismo epistemologico: alternative praghesi", in *Le ragioni del conoscere e dell'agire. Scritti in onore di Rosaria Egidi*, ed. by R. M. Calcaterra (Franco Angeli, Milano 2006), 190-204. Eng. trans. by the author, copy-edited (with permission) by K.R. Westphal.

philinq IX, 1-2021, pp. 79-94
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.358

enabled him to leave the United States and go to Europe in 1932. After having worked with Carnap, in the second semester of 1934 he moved to Warsaw. Here he came into contact with the Polish logicians and in particular Alfred Tarski, with whom in the meantime Carnap himself had had a positive exchange of ideas (see Quine 1986: 10-13; and Carnap, 1963a: 30). In the course of 1934, Carnap, Tarski and Quine had no opportunity to discuss their ideas all three together. I do think, however, that the contacts which they had separately with each other during that year must be ideally connected to the those they had a few years later, in the academic year 1940-41, when all three of them met at Harvard. In fact, what clearly emerges from their Harvard discussions (in which Bertrand Russell too participated actively) was the clear-cut disagreement between Carnap, on the one hand, and Tarski and Quine, on the other, about the possibility of maintaining a "sharp […] distinction between logical and factual truth" (Carnap 1963a: 35-36). However, if we consider the developments of Carnap's and Quine's thought more closely, we can see that the main ideas which characterize the theoretical gap between their positions had already emerged in the Prague period.

In the years immediately preceding Carnap's move to Prague, he had been working on *Der Logische Aufbau der Welt.* This book was published in 1928, but was the fruit of a long preparation begun in the early 1920s, when Russell's and Wittgenstein's influence had not yet fully shown its effect. As we can see from Frank's own words, the very ideas Carnap was working on in these years provide one of the main reasons why Schlick wanted him in Vienna. In these ideas the Vienna philosophers saw the attempt to realize that empiristic synthesis between Kantism, Mach's sensism, Poincarè's conventionalism and the new mathematical logic which was to underwrite a renewed defense of scientific rationality, capable of opposing the well-known thesis of the bankruptcy of science.[2]

The positions contained in the *Aufbau*, though, were not immune from criticism, also from within the Vienna Circle itself, in particular from the physicalist Otto Neurath. The debate about protocols led Carnap to abandon

[2] Frank writes: "According to Mach the general principles of science are abbreviated economical descriptions of observed facts; according to Poincaré they are free creations of the human mind that do not tell anything about observed facts. The attempt to integrate the two conceptions into one coherent system was the origin of what was later called Logical Empiricism. […] Carnap gave the new philosophy its 'classical' shape. [… In] his book *The Logical Structure of the World* [(*Der logische Aufbau der Welt*) …] the integration of Mach and Poincaré was actually [achieved] in a coherent system of conspicuous logical simplicity. Our Viennese group saw in Carnap's work the synthesis that we had advocated for many years" (Frank 1949, 11-12, 33). On the importance of this evidence given by Frank in understanding Logical Empiricism, see the essays collected in Parrini (2002), in particular, Chapters 1 and 6.

the kind of phenomenalistic reductionism which is one main characteristic of the *Aufbau* and turn instead in a more strongly conventionalist and anti-foundationalist direction regarding both the nature of the empirical basis and the epistemological status of logic and mathematics. The clearest expression of such a position can be found in the *Logische Syntax der Sprache*, a book which Carnap had first conceived also under the influence of some suggestions offered to him by Tarski in Vienna in February 1930[3] and which he published a few years later with a *Vorwort* dated "Prag, im Mai 1934".

In the *Logische Syntax der Sprache* Carnap deals both with specific questions of philosophy of logic and mathematics and questions of a more general kind. I will now focus on three theses contained in this work.

The first thesis regards logical-mathematical conventionalism. It can be summarized in the famous *Toleranzprinzip*: "In logic, there are no morals (*In der Logik gibt es keine Moral*)" (Carnap 1934/1937: §17). Logical and mathematical *a priori* truths depend only on *L*-rules, in other words on the conventions (in effect: implicit definitions) that fix the meaning of logical-mathematical symbols. In this way, logical-analytical truths, called *L*-truths or *L*-valid propositions, are distinguished from *P*-truths, *i.e.* those truths that depend on postulates (known as *P*-rules) at the basis of the theories of empirical sciences, physics in particular. The so called "linguistic doctrine of the *a priori*" is here presented in a conventionalistic form, since referring to the method of implicit definitions allows us to eliminate any reference to forms of *a priori* intuition in characterizing either logical-mathematical truths or other scientific general principles, such as the causal principle.

The second thesis is linked to the distinction between formal and material language. According to Carnap, the "*material mode of speech*" is "*a transposed mode of speech*" and the "*translatability into the formal mode of speech constitutes the touchstone for all philosophical sentences* or, more generally, for all sentences which do not belong to the language of any one of the empirical sciences" (Carnap 1934/1937: §80). The task of philosophical analysis consists in eliminating philosophical problems by translating them into the formal language, or reinterpreting them as questions of a practical nature regarding the form of the language we intend to adopt.

Such an idea (which, in substance, can already be found in the *Aufbau*) will be fully expressed in his 1950 essay, "Empiricism, Semantics and Ontology". In this essay, the distinction between *L*-rules and *P*-rules – which he had improved

[3] See Carnap (1963a: 30). Here Carnap recalls that his disagreement with Tarski about the "difference between logical and factual statements" had already emerged in those years, because Tarski "maintained that the distinction was only a matter of degree."

in his works on semantics subsequent to the *Logische Syntax der Sprache* (from the *Introduction to Semantics* to the numerous Appendixes to the second, 1956 edition of *Meaning and Necessity*[4]) – is connected to the distinction between *internal* and *external* questions. Answers to internal questions are given within a linguistic framework already accepted in accord with the rules that characterize it. These answers depend on the particular nature of the internal questions involved and may be empirical answers or logical-analytical answers. External questions, instead, regard the problem of which linguistic frameworks we should accept. The answers given to them are of a substantially pragmatic nature, although decisions to accept or reject a framework also consider the empirical and theoretical factors we have at our disposal (see Carnap 1956, esp.: 206-209).

The third thesis of the *Logische Syntax der Sprache* on which I want to focus my attention provides the general philosophical background to all the others. This thesis consists in the linguistic-syntactic form that the anti-absolutistic component of the *wissenschaftliche Weltauffassung* takes in Carnap's views. Such a component has been accurately described by Frank in his brief reconstruction of the origins and development of Logical Empiricism,[5] where it is traced back to the anti-metaphysical spirit that animates Mach's treatment of physics, in particular mechanics:

> Mach analyzed the fundamental concepts of nineteenth-century physics, such as mass and force, and made clear that all statements containing these words can be interpreted as statements about sense observations. […] Nonetheless, Mach had no special bias against the mechanistic terminology that would imbue him with a particularly antimaterialistic tendency. He tried to debunk all types of auxiliary concept in so far as they pretended to describe ontological realities or metaphysical entities.[6]

Carnap links this idea to the verification principle and the linguistic conception of the *a priori* and supports the thesis that we can discuss questions relating

[4] See Carnap (1956, 1st ed.: 1947). Among the numerous essays reprinted in the appendix to the second edition of *Meaning and Necessity*, the one published in 1955 ("On Some Concepts of Pragmatics") is particularly significant (Carnap 1956: 248-250). While answering R. Chisholm's objections, Carnap acknowledges the link between intentional notions (such as belief) and intensional notions (such as meaning). Moreover, continuing the discourse begun in the 1952 essay "Meaning Postulates" (Carnap 1956: 222-229), he states the theoretical nature of semantic and pragmatic concepts; see Parrini (1976, esp.: 97-116), and Creath (1990: 1-43). Creath's analysis, though, gets only so far as Carnap's essay, "Meaning and Synonymy in Natural Languages" (Carnap 1956: 233-247), which was written between "Meaning Postulates" and "On Some Concepts of Pragmatics"; see Creath (1990, esp.: 34-38).

[5] See Frank (1949: 1-53), "Introduction – Historical Background".

[6] Frank (1949: 17-18). See Parrini (1998: 13-16). In the paper presented in Prague, Thomas Uebel underlined the importance of this Machian aspect of Frank's thought by speaking of the tendency to the metaphysical hypostatization or absolutization of scientific concepts as a form of "petrification".

to existence and truth in a meaningful way only relative to a linguistic framework previously specified: asking questions of truth and existence has only an internal empirical-cognitive meaning. In the Prague years, the framework is conceived as a structure of a purely syntactic nature. According to Carnap,

> the use of the material mode of speech gives rise to obscurity by employing absolute concepts in place of the syntactical concepts which are relative to language […]. The use of the *material mode of speech* leads […] to a *disregard of the relativity to language of philosophical sentences*; it is responsible for an *erroneous conception of philosophical sentences as absolute.* (Carnap 1934/1937, §80)

Just when discussing analyticity in the characterization given to it by Frege, Carnap quotes Walter Dubislav in order to state the relative nature of this notion. We can speak of analytical propositions only relatively "to a particular system of assumptions and methods of reasoning (primitive sentences and rules of inference), that is to say, in our terminology, to a particular language."[7] Such a thesis is connected to the conventionalism explicated in the *Logische Syntax* and the criticism developed in this same book opposing "Wittgenstein's absolutist conception of language, which leaves out the conventional factor in language-construction."[8] It is important to notice that Carnap does not simply say he is interested in elaborating a relativized conception of analyticity, or more precisely, *L*-validity. He states a stronger thesis: to the notion of analyticity, as well as to other similar notions, we should only ascribe a relative validity.

All these ideas had a profound influence on Quine, who more than once has acknowledged his intellectual debt.[9] His thought, though, will culminate in a theory that – as Richard Creath (1991) rightly pointed out – comprises an epistemological project which is alternative to Carnap's. Unlike Carnap, Quine was not mainly interested in developing a model of epistemic justification of our assertions in which, side by side to experience, an essential role is played by conventions and meanings and not by forms of *a priori* intuition and *a priori* concepts or principles. Quine's epistemological project regards primarily the transformation of the complex of our beliefs and convictions and is mainly linked (though not exclusively) to the idea of naturalization. Quine takes Neurath's well-known metaphor of the sailors very seriously and tries to describe the process by which we try to improve the system of

[7] Carnap (1934/1937: §14). Here Carnap refers, as well as to Kant, Frege and Wittgenstein, to Walter Dubislav's essay, *Über die sogenannten analytischen und synthetischen Urteile* (Berlin, 1926).

[8] Carnap (1934/1937: §52); see also §67, where Carnap accuses Wittgenstein of overlooking "the fact that there is a multiplicity of possible languages" and talking "continually of 'the' language".

[9] See, for example, Carnap & Quine (1990: 463-466).

beliefs in which from the very beginning we have been immersed by calling into question considerations of a global nature governed by the principles of empirical adequacy, simplicity and conservation. In this way, in Quine's epistemology (or at least, in Quine's *desiderata*), no role is played by forms of *a priori* intuition (as also in Carnap's conception too), but also none is played by conventions, meanings or distinctions between the *a priori* and the *a posteriori* and the analytic and the synthetic, of whatever nature they may be (relativized or non-relativized).

Quine presented such a conception – a very problematic one, especially in its naturalized version – in "Two Dogmas of Empiricism" and in the other essays published with it in *From a Logical Point of View* (1953). I think, though, that he had already set the essential premises both of his criticism of Carnap and of his own alternative epistemological project in "Truth by Convention," an essay published in 1936 but which he completed in 1935, one year after his Prague period. In "Truth by Convention", Quine critically analyses the idea that logical-mathematical truths depend upon conventions concerning the meaning or linguistic use of logical constants. Yet he does not go so far as refuting the analytic/synthetic and the *a priori/a posteriori* distinctions. Despite this, in his analysis of logical conventionalism we can already see three fundamental aspects of his future position.

The first aspect is his naturalistic behaviorism. Quine points out that, if one wished to, it would be possible to apply the method of implicit definition not only to logic and mathematics, but also to the "so called empirical sciences" (Quine 1936: 100) extending the conventionalistic thesis to them too. If we do not do this, it is because by asserting the conventionality of logical-mathematical truths, but not the conventionality of the empirical truths, we want to account for the fact that "the former are *a priori*, the latter are *a posteriori*; the former have 'the character of an inward necessity', in Kant's phrase, the latter do not" (Quine 1936: 102). In discussing this point Quine states that it is possible to look at the contrast between the two types of truth from a strictly behavioristic point of view, "and without reference to a metaphysical system," "as a contrast between more and less firmly accepted statements" which "obtain antecedently to any *post facto* fashioning of conventions":

> there are statements which we choose to surrender last, if at all, in the course of revamping our sciences in the face of new discoveries, and among these there are some which we will not surrender at all, so basic are they to our whole conceptual scheme. Among the latter are to be counted the so-called truths of logic and mathematics, regardless of what further we might have to say of their status in the course of a subsequent sophisticated philosophy. (Quine 1936: 102)

The second aspect of "Truth by Convention" I want to underline proves that logical-mathematical conventionalism is invalidated by an infinite regress. Logical truths are infinite in number; so they cannot be singled out individually. In order to indicate them, it is necessary to advert to general conventions; but to apply general conventions to individual cases, we already need logic at a meta-theoretical level: if logic "is to proceed *mediately* from conventions, logic is needed for inferring logic from the conventions" (Quine 1936: 104). Quine's argument – which I do not need to expound here in full detail – is substantially the argument identified by Lewis Carroll in his 1895 essay "What the Tortoise Said to Achilles." Such an argument was mentioned again in Quine's (1954/1963) essay, "Carnap and Logical Truth",[10] and his allied attempts to defend the thesis of the conventionality of logic up to today. For example, it constitutes one of the major obstacles that must be met by the epistemic conception recently taken again into consideration by Paul Boghossian.[11]

The third aspect characterizing Quine's criticism of logical conventionalism is particularly relevant, since it anticipates the thesis according to which no genuine *explicandum* corresponds to the analytic/synthetic dichotomy. While commenting on the question of the infinite regress, Quine is willing to concede that we can deal with it by maintaining that the conventions necessary to produce logical and mathematical truths "are *observed* from the start, and that logic and mathematics thereby become conventional":

> It may be held that we can adopt conventions through behavior, without first announcing them in words; and that we can return and formulate our conventions verbally afterwards, if we choose, when a full language is at our disposal. (Quine 1936: 105-106)

Straight afterward, though, Quine adds that this kind of defense risks depriving the notion of convention of any recognizable content. In such a case – Quine says – "it is not clear wherein an adoption of the conventions, antecedently to their formulation, consists; such behavior is difficult to distinguish from that in which conventions are disregarded":

> In dropping the attributes of deliberateness and explicitness from the notion of linguistic convention we risk depriving the latter of any explanatory force and reducing it to an idle label. We may wonder what one adds to the bare statement that the truths of logic and mathematics are *a priori*, or to the still barer behavioristic statement that they are firmly accepted, when he characterizes them as true by convention in such a sense. (Quine 1936: 105-106)

[10] See Quine (1954: 115); on "Carnap and Logical Truth", see Creath (2003).

[11] See, for example, Boghossian (2003).

So – Quine concludes – "as to […] the thesis that mathematics and logic proceed wholly from linguistic conventions, only further clarification can assure us that this asserts anything at all" (Quine 1936: 105-106).

With this third point Quine not only posited one of the building blocks for his future criticism of the two dogmas of empiricism and of Carnap's conception of "semantic ascent"; he also posited one of the major problems which worried Carnap in his defense of the notion of analyticity: the problem of the *explicandum* that such a notion should account for. This problem presents two aspects: a semantic-pragmatic aspect which pertains to the philosophy of language,[12] and an epistemological aspect that pertains to the general theory of knowledge and to philosophy of science. Taken in its epistemological meaning, it is just this problem that will lead Carl Gustav Hempel to side with Quine regarding the possibility of maintaining the analytic/synthetic distinction.

The reason why Hempel's path crossed Quine's is linked to a technical question concerning the formulation of the Standard Conception of Scientific Theories. In the course of the liberalization of empiricism, Hempel had shown that the method of the "so-called bilateral reduction sentence" used by Carnap to provide an empirical interpretation of dispositional and theoretical terms made it difficult to keep separate "the specification of meanings and the description of facts" (Hempel 1963: 686, 691). Under the stimulus of such critical observations, Carnap managed to devise a very ingenious and complex solution based on the use of the "Ramsey sentence" which allowed a reconstruction of theories in which the analytical components were distinguished from those which are synthetic.

In the essay published in Schilpp's volume on Carnap, Hempel acknowledged the success of Carnap's solution from a strictly technical point of view, but objects to its epistemological relevance. He maintains that the "new procedure" devised by Carnap "gives rise […] to the question as to the meaning and the rationale of the distinction that is made here between meaning postulates and empirical postulates" (Hempel 1963: 705). Referring to Quine's criticism of reductionism, Hempel states again that in science there are no assertions totally devoid of empirical content, the truth value of which cannot be revised in the light of future experiences. Thus, he deems

> questionable […] whether there is any aspect of scientific method or of scientific knowledge that would constitute an explicandum for the analytic-synthetic dichotomy in regard to the statements of empirical science. (Hempel 1963: 705)

[12] For the semantic-pragmatic aspect, see above, note 4 and the references given there.

It seems to me significant that in his comment on Hempel's contribution, Carnap (1963c) does not answer this problem. Actually, he could rightly have thought that he had already explicitly dealt with it in his comment on Quine's essay, "Carnap and Logical Truth". In fact, there he had specified the following three points: (*i*) "the concept of analytic statement" he had adopted "as an explicandum is not adequately characterized as 'held true come what may'" (Carnap 1963b: 921); (*ii*) in scientific developments it is opportune to distinguish between readjustments in the attribution and/or change of truth values assigned to statements within a given language and readjustments of a "revolutionary" kind which render the change of the language (linguistic form) of reference necessary; (*iii*) his "concept of analyticity as an explicandum has nothing to do with such a transition"; it refers "in each case to just one language":

> That a certain sentence *S* is analytic in [a particular language *L*] means only something about the status of *S* within the language [*L*]; as has often been said, it means that the truth of *S* in [*L*] is based on the meanings in [*L*] of the terms occurring in *S*. (Carnap 1963b: 921)

Only in the last few decades – thanks to studies which have considerably deepened our knowledge of both the historical development of Logical Empiricism and the relations between neo-empiricistic conceptions (Carnap's in particular) and the ideas of the so-called New Philosophy of Science (Kuhn's in particular) – has it become possible to fully understand the sense of the position taken by Carnap in the controversy on analyticity. In contrast to what was initially believed, between Carnap's and Kuhn's conceptions there is not only a contraposition. On the one hand, it is certainly true that Kuhn's ideas led to the crisis of a 'vertical conception' of science, characterized by the dualism between theoretical language and observational language. On the other hand, it is equally true that such ideas added extra value to the thesis – characteristic also of Carnap's epistemology with its distinction between internal and external questions – according to which we cannot understand the structure of and changes within science without taking into consideration the presuppositions that provide the framework for scientific activity. While considering the holistic conception of the theory/experience relation still to be valid, we cannot consider adequate a vision of science which puts all the expressive components of scientific discourse on the same level, without setting any distinction between those components which depend upon experience directly and those which depend on it only indirectly and play a presuppositional role.

In Carnap's epistemology the rules, or meaning postulates, and the analytic statements depending upon them, play a role analogous to the role played by the so called "paradigmatic propositions" in Kuhn's contraposition between

normal and revolutionary science. Together with manuals and exemplars (or paradigms in the central sense of the term), such propositions are an essential component of the disciplinary matrix and are statements that can be considered neither as necessarily valid,

nor empirical in the usual sense exactly because they are protected from straightforward empirical refutation […]. They constitute an epistemically distinct class in that they do not fit the traditional division of all propositions into *a priori* and empirical. Rather they are propositions which are accepted as a result of scientific experience but which come to have a constitutive role in the structure of scientific thought.[13]

Recently it has been pointed out that an analogous problem had already been posited (though only with reference to logical principles) in a letter Goodman wrote to Quine in the early 1950s (see Creath 1991: 380-381). It should be noticed, though, that in the *Logische Syntax* Carnap had already set the premises of the answer he later gave Hempel and Quine in the 1960s. In §82 ("Physical Language") of that book, he says that "either *L*-rules alone, or *L*-rules and *P*-rules, can be laid down as transformation rules of the physical language" (Carnap 1934/1937: §82). Furthermore – well before Quine's revival of Duhem's thesis in "Two Dogmas of Empiricism" – on the basis of the results of the polemic about protocols, he states the holistic character of the experimental control and the revisability in principle of any statement, in other words not only of protocol statements and *P*-rules, but also of *L*-rules. In fact, Carnap declares that the empirical test of hypotheses and theories is relative to other hypotheses and theories:

the test applies, at bottom, not to a single hypothesis but to the whole system of physics as a system of hypotheses (Duhem, Poincaré). No rule of the physical language is definitive; all rules are laid down with the reservation that theory may be altered as soon as it seems expedient to do so. This applies not only to the *P*-rules but also to the *L*-rules, including those of mathematics. In this respect there are only differences in degree; certain rules are more difficult to renounce than others. (Carnap 1934/1937: §82)

It is important to notice that in the *Logische Syntax* Carnap maintains this point drawing indifferently from both Poincaré and Duhem. As in Frank's Introduction to *Modern Science and Its Philosophy* (1941), Carnap too does not seem to be aware of the relevant differences between Poincaré's position and Duhem's, differences that Duhem himself had strongly underlined in some (for too long neglected) pages of his *Thèorie physique*.[14] Here I must set aside this aspect of

[13] Brown (1979: 105); see also Kuhn's concise elucidations in Kuhn (1983: 566-567).

[14] See Frank (1949); Frank's pages 15-16 on Duhem are particularly relevant. Frank underlines Duhem's holistic conception of experimental control, but does not take into consideration his criticism

the question and focus instead on the most relevant traits of the conception of the theory/experience relation proposed in the passage mentioned above of the *Logische Syntax*.

There is no doubt that in supporting the global dependence upon experience of the system of our assertions and the revisability in principle of each of its components, Carnap was proposing a thesis that later played a fundamental role both in Quine's conception of an empiricism without dogmas and in the criticism directed by Quine to Carnap himself about the possibility of maintaining the analytic/synthetic, *a priori/a posteriori* and internal/external distinctions. It is also unquestionable, though, that Carnap supported these ideas in the context of a very precise thesis: the above mentioned relativistic thesis, according to which the questions regarding truth and existence can be raised only within a linguistic framework previously established. As I have tried to show elsewhere, this thesis was not undermined either by Quine's subsequent criticisms of the two dogmas of empiricism nor by the abandonment of the verification principle[15]; on the contrary, it was even reinforced by what Kuhn stated about scientific revolutions. In this way, Carnap stepped forward towards formulating that conception of the relativized *a priori* that some recent interpreters (including myself and Michael Friedman[16]) have considered as one of the most characteristic points of the Neo-empiristic conception of the theory/experience relation.

In fact, with his defense of analyticity Carnap aligned himself with the work in which the idea of the relativized *a priori* had been outlined for the first time. I refer to the 1920 book *Relativitätstheorie und Erkenntis a priori* in which Reichenbach stated the necessity to maintain the idea of a constitutive *a priori* endowed with a double nature (as later Kuhn's so-called "paradigmatic propositions" will be): such an *a priori* is subject to historical changes (in other words, it is not eternally valid) and it is not absolutely independent of experience. In fact, the main characteristic of Reichenbach's coördinative or constitutive principles is that

of Poincaré on the language/theory relationship. On the importance of this topic for the interpretation of Logical Empiricism, see Parrini (2002), esp. Chapters 1, 6. Duhem's holism is also not discussed in the collection of essays edited by P. Frank (1961), *The Validation of Scientific Theories* (Collier Books, New York). (The papers published in this book were first presented at the annual meeting of the American Association for the Advancement of Science, Boston, Massachusetts, December 1953.)

[15] Parrini (1994, esp. 267-274), and Parrini (2002: chapter 10).

[16] See: Parrini (1976, esp.: 264-290); Parrini (2002: chapters 6, 7, 10); Friedman (1999), (2001), esp. Part Two, "Fruits of Discussion", "The Relativized *A Priori*", 79, *n.* 9. As a scrupulous and intellectually fair scholar once said: "dates are clear."

their validity does not depend only upon the judgment of particular experiences, but also upon the possibility of the whole system of knowledge: this is the sense of the *a priori*. The fact that we can describe reality by means of metric relations among four coördinates is as valid as the totality of physics; only the special form of these rules has become a problem of empirical physics. This principle is the basis for the conceptual construction of physical reality. *Every* physical experience ever made has confirmed this principle. This result does not exclude the possibility that some day experiences will occur that will necessitate another successive approximation – then physics again will have to change its concept of object and presuppose new principles for knowledge. '*A priori*' means '*before* knowledge', but not 'for all time' and not 'independent of experience'. (Reichenbach 1920: 104-105)

In passing from Reichenbach's to Carnap's conception, the relativized *a priori* undergoes a significant change that has often been overlooked: The coördinative assumptions mentioned by Reichenbach loose their theoretical-synthetic nature to become linguistic conventions in Poincaré's sense (as Schlick had already stated). Such assumptions cease being constitutive of objects and become constitutive of meanings. Just on this point we can appreciate how failing to see the differences between Poincaré's position and Duhem's has been relevant within the history of Logical Empiricism. In fact, one of the main criticisms addressed to Poincaré by Duhem aimed precisely at establishing that those aspects of subjectivity present in scientific discourse cannot be reduced (as Poincaré claimed) to the linguistic component of such discourse. Putting aside this question which I have discussed elsewhere (Parrini 2002: Chapters 1, 6, 10), here I wish to draw the attention to the fact that the position expounded by Carnap in the *Logische Syntax* implied a way of looking at analyticity and the *a priori* radically different from, and alternative to, the position Quine maintained until the end of his career in the 1990s.

In one of the Replies contained in the Schilpp volume, Quine says:

I now perceive that the philosophically important question about analyticity and the linguistic doctrine of logical truth is *not* how to explicate them; it is the question rather of their relevance to epistemology. The second dogma of empiricism, to the effect that each empirically meaningful sentence has an empirical content of its own, was cited in "Two Dogmas" merely as encouraging false confidence in the notion of analyticity; but now I would say further that the second dogma creates a need for analyticity as a key notion of epistemology, and that the need lapses when we heed Duhem and set the second dogma aside. (Quine 1986a: 207)

This passage is important for a number of reasons, not least because it testifies to one of Quine's many oscillations when trying to state the reasons of his re-

jection of analyticity.[17] In this case, the quotation helps us understand that one of the essential points to consider when discussing the *a priori/a posteriori* and analytic/synthetic distinctions is their relevance to epistemology. This remains true also after reductionism has been rejected in favor of epistemological holism and (still more significant) rightly so, due to reasons strictly linked to a holistic conception of the theory/experience relation.

It may be that the negation of analyticity must be considered necessary to the development of holism according to Quine's model (although I do not agree with this consideration, at least if we remain on a strictly epistemological level[18]); but if we look at the question from the point of view of epistemic justification, it does seem that the holistic conception of the theory/experience relation as such requires re-evaluation of the relativized *a priori*. Unlike what Quine maintained, we can renounce reductionism in favor of Duhem's holism (as Carnap had done since the time of the *Logische Syntax*) without depriving analyticity of its key role in epistemology.

The most recent discussions of analyticity and the *a priori* have brought to light merits and demerits of both Quine's and Carnap's conceptions. Just these discussions have allowed us to understand that distinctions such as the analytic/synthetic and *a priori/a posteriori* distinctions cannot be denied or accepted when considered individually in isolation, *i.e.* without taking into consideration the answers we give to other philosophical questions. Such distinctions must be accepted or rejected as integral parts of distinctive general epistemological theories, similar to those alternative conceptions which began to emerge in Prague in the first half of the 1930s and which still are pivotal points in contemporary epistemological debate.

References

Boghossian, Paul, 2003, "Epistemic Analyticity: A Defense," in Glock, H.-J., Glüer, K., & Keil, G., eds., *Fifty Years of Quine's "Two Dogmas*," in *Grazer Philosophische Studien*, 66: 15-35.

Brown, Harold, 1979, *Perception, Theory and Commitment. The New Philosophy of Science*, The University of Chicago Press, Phoenix edition, Chicago; 1st ed. 1977.

[17] For more detailed analysis of this issue see Parrini (2018: esp. §§2-4). [This companion piece follows in this issue. – *Ed. note*.]

[18] I think that we cannot justify the statement that the dogma of analyticity and the dogma of reductionism "are, indeed, at root identical," affirmed in "Two Dogmas" (Quine 1951: 41). For this reason I think it is possible to reject the second dogma while maintaining the first. We can preserve analyticity even when accepting the revision argument: it is enough to say – as Carnap said – that the revision of the truth value of certain sentences involves a change in language. Deprived of its semantic-pragmatic objections, Quine's position on analyticity loses much of its appeal and strength.

Carnap, Rudolf, 1934, *Logische Syntax der Sprache*, Springer Verlag, Wien; cited by § number.

Carnap, Rudolf, 1937, *The Logical Syntax of Language*, Routledge & Kegan Paul, London, 3rd ed. 1951; cited by § number.

Carnap, Rudolf, 1950, "Empiricism, Semantics, and Ontology", in *Revue Internationale de Philosophie*, 4: 20-40; rev. ed. in: *idem* (1956): 205-221.

Carnap, Rudolf, 1956, *Meaning and Necessity. A Study in Semantics and Modal Logic*, 2nd rev. ed., University of Chicago Press, Chicago; 1st ed. 1947.

Carnap, Rudolf, 1963a, "Carnap's Intellectual Autobiography", in Schilpp, P.A., ed., *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, La Salle ILL: 1-84.

Carnap, Rudolf, 1963b, "W.V. Quine on Logical Truth", in Schilpp, P.A., ed., *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, La Salle ILL: 915-922.

Carnap, Rudolf, 1963c, "Carl G. Hempel on Scientific Theories", in Schilpp, P.A., ed., *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, La Salle ILL: 958-966.

Carnap, Rudolf, & Quine, W.V.O., 1990, Creath, Richard, ed., *Dear Carnap, Dear Van. The Quine-Carnap Correspondence and Related Work*, University of California Press, Berkeley & Los Angeles.

Carroll, Lewis, 1895, "What the Tortoise Said to Achilles", in *Mind* 4, 14: 278-280.

Creath, Richard, 1990, "Introduction", in Carnap & Quine (1990): 1-43.

Creath, Richard, 1991, "Every Dogma Has Its Day", in *Erkenntnis*, 35: 347-389.

Creath, Richard, 2003, "The Linguistic Doctrine and Conventionality: The Main Argument in 'Carnap and Logical Truth'", in Hardcastle, G.L., & Richardson, A.W., eds., *Logical Empiricism in North America*; *Minnesota Studies in the Philosophy of Science* XVIII, University of Minnesota Press, Minneapolis: 234-256.

Frank, Paul, 1949, *Modern Science and its Philosophy*, Harvard University Press, Cambridge MA; 1st ed. 1941.

Friedman, Michael, 1999, *Reconsidering Logical Positivism*, Cambridge University Press, Cambridge.

Friedman, Michael, 2001, *Dynamics of Reason*, CSLI Publications, Stanford, CAL.

Hempel, Carl, 1963, "Implications of Carnap's Work for the Philosophy of Science", in Schilpp, P.A., ed., *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, La Salle ILL: 685-709.

Kuhn, Thomas, 1983, "Rationality and Theory Choice", in *The Journal of Philosophy*, 80, 10: 563-570.

Parrini, Paolo, 1976, *Linguaggio e teoria. Due saggi di analisi filosofica*, La Nuova Italia, Firenze.

Parrini, Paolo, 1994, "With Carnap Beyond Carnap. Metaphysics, Science, and the Realism/Instrumentalism Controversy", in Salmon, Wesley, & Wolters, Gereon, eds., *Logic, Language, and the Structure of Scientific Theories*, Proceedings of the Carnap-

Reichenbach Centennial, University of Konstanz, 21-24 May 1991; University of Pittsburgh Press/Universitätsverlag Konstanz, Pittsburgh/Konstanz: 255-277.

Parrini, Paolo, 1998, *Knowledge and Reality. An Essay in Positive Philosophy*, Kluwer, Dordrecht; Eng. ed. of *Conoscenza e realtà. Saggio di filosofia positiva*, Laterza, Roma-Bari 1995.

Parrini, Paolo, 2002, *L'empirismo logico. Aspetti storici e prospettive teoriche*, Carocci, Roma.

Parrini, Paolo, 2018, "Quine su analiticità e olismo. Una valutazione critica in dialogo con Sandro Nannini", in Lumer, C., & Romano, G., eds., *Dalla filosofia dell'azione alla filosofia della mente. Riflessioni in onore di Sandro Nannini*, Corisco Edizioni, Roma-Messina: 95-116. [Eng. trans. follows below – *Ed. note*.]

Quine, W.V.O., 1936, "Truth by Convention", in Lee, O.H., ed., *Philosophical Essays for A.N. Whitehead*, Longmans, New York; rpt. in: W.V.O. Quine, *The Ways of Paradox and Other Essays*, rev. & enlarged edition, Harvard University Press, Cambridge MA: 77-106.

Quine, W.V.O., 1951, "Two Dogmas of Empiricism", in *Philosophical Review*, 60: 20-43; rpt. in: *idem*, *From a Logical Point of View. Logico-philosophical Essays*, 2nd rev. ed., Harper & Row, New York 1961: 20-46.

Quine, W.V.O., 1954, "Carnap and Logical Truth", rpt. in: *idem*, *The Ways of Paradox and Other Essays*, rev. & enlarged ed., Harvard University Press, Cambridge MA: 107-132.

Quine, W.V.O., 1986, "Autobiography of W. V. Quine", in Hahn, L.E., & Schilpp, P.A., eds., *The Philosophy of W.V. Quine*, Open Court, La Salle ILL: 1-46.

Quine, W.V.O., 1986a, "Reply to Geoffrey Hellman", in Hahn, L.E., & Schilpp, P.A., eds., *The Philosophy of W. V. Quine*, Open Court, La Salle ILL: 206-208.

Reichenbach, Hans, 1920, *Relativitätstheorie und Erkenntnis apriori*, Springer, Berlin; Eng. ed.: Reichenbach, Maria, ed., *The Theory of Relativity and A Priori Knowledge*, University of California Press, Berkeley & Los Angeles 1965.

Quine on Analyticity and Holism: A critical appraisal in dialogue with Sandro Nannini[1]

Paolo Parrini†

Abstract: The first four sections evaluate Quine's thesis that the two dogmas of empiricism (analyticity and reductionism) are at root identical. In particular, a full compatibility is developed and defended between epistemological, anti-reductionist holism and both the analytic/synthetic and *a priori/a posteriori* distinctions. According to the view defended here, understanding the relations between theory and experience requires not the rejection of such distinctions, but rather their relativization. In the fifth and final section, the importance of such distinctions is shown regarding epistemological analysis and discussions of the relations between science and philosophy.

1. I am particularly pleased to participate in this initiative in honor of Sandro, both because of our old friendship and the passionate and lively philosophical discussions we had year after year.

Such discussions have chiefly concerned two themes: realism and naturalism. Sandro is in favor of what I would call metaphysical realism, whereas I am in favor of empirical realism.[2] Sandro supports a naturalism explicitly deriving from Quine, whereas I have some methodological misgivings about the turn taken by the debate on the *mind-body problem*[3] in the last decades, let alone about naturalism. Even though, in principle, I have no objection to the programs of naturalization, I still think that epistemology and phenomenology posit prob-

[1] English version of *Quine su analiticità e olismo. Una valutazione critica in dialogo con Sandro Nannini*, in C. Lumer & G. Romano, eds., *Dalla filosofia dell'azione alla filosofia della mente. Riflessioni in onore di Sandro Nannini*, Corisco Edizioni, Roma-Messina 2018. [Author's translation, copy-edited by permission by K. R. Westphal.]

[2] See Nannini and Parrini in Lanfredini & Peruzzi (2013, 2016), respectively vol. 1: 113-127, vol. 2: 61-88, in particular, 75-77. I must add that the reasons why Sandro does not share my position on realism have nothing to do with some recent singular evaluations of it. In my opinion, certain appraisals fail to consider the various aspects involved in the *Realismusfrage*. For similar reasons, it seems to me that also other more elaborate criticisms rest on misunderstandings of my ideas or reduce to clear forms of *begging the question* (see below *n.* 4).

[3] On this point, I share many observations contained in Westphal (2016).

philinq IX, 1-2021, pp. 95-112
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.359

lems which, at least so far, do not seem dissolved, nor satisfactorily solved, by naturalistic conceptions bordering on physicalism (Nannini 2015, Parrini 2015).

Today I would like to speak about a third theme that so far has remained in the background, though likely lying at the origin of our different positions on both realism and naturalism. I refer to our attitude towards the way in which Quine developed his holism by rejecting the analytic/synthetic distinction and the *a priori/a posteriori* distinction which he considered (as do Logical Empiricists) co-extensive with the analytic/synthetic distinction. The forerunner of our divergence can be found in one of Sandro's books written in the form of a dialogue: *La Nottola di Minerva. Storie e dialoghi fantastici sulla filosofia della mente* (2008).

In the copy he kindly gave me, Sandro wrote that one of the characters of such a dialogue, namely the *Analitico Primo* (who above all seems to reflect the standard Neo-empiristic conception), "owes a lot" to me. To tell the truth, my epistemological position (which centrally highlights, *inter alia*, an interactive theory of knowledge, the reticular model, the idea of an open-texture rationality and the negation of transcendental principles of knowledge of a Kantian kind) is rather a conception of a Neurathian, post-Neo-empiristic and post-Quinean sort, far closer to the perspectives advocated by another character of Sandro's book, the *Straniero*, than to those of the old Logical Empiricism. Since many of the theses supported by the *Straniero* are very dear to Sandro, this means that, apart from realism and naturalism, our ideas appear to converge considerably.

Of course, there is a link between me and the *Analitico Primo*, but this link only grasps the fact that, although in a way and in a context deeply modified, I have tried to keep a role for some Neo-empiristic ideas criticized by Quine. In particular, although I no longer accept a fundamental Neo-positivistic principle such as the verification principle, I have defended a modified version of the *a priori/a posteriori* and analytic/synthetic distinctions which were integral parts of the Neo-empiristic package. Here I wish to show in which way such distinctions can coexist with a firm assent to a holistic conception of the theory/experience relation, such as that supported by Quine (though with an odd oscillation I will mention shortly). I shall try to explain, first, the reasons why I think that holism requires not a rejection, but a relativization of the analytic/synthetic and *a priori/a posteriori* distinctions, with a weak (not a strong) negation of the Kantian synthetic *a priori*; second, the reasons why I believe that such distinctions are *philosophically* crucial in order to answer problems regarding objectivity, truth and realism.[4]

[4] The complexity of the problem of realism referred to in note 2 depends upon this: The answers to the questions of objectivity, reality and truth must consider many conflicting elements among

To avoid misunderstandings, I specify in advance that when in the next pages I clarify my position with respect to the assertions of the *Straniero* and the *Psicologa* (another character of *La Nottola di Minerva*), I do not intend in the least to criticize Sandro, whose detailed convictions about the relation between holism and the *a priori* I do not know. I use his dialogue and characters only because, though he does not entirely accept "the classic scientific realism of a Quinean origin" (Nannini 2013: 120), he appears to be so near to a kind of naturalized epistemology as to believe "in the *collapse* of the distinction between analytical judgements and synthetical judgements" (Nannini 2013: 123, emphasis added). For this reason, I suspect that some assertions of the *Straniero* and the *Psicologa* may be the clue to a certain divergence between the two of us, not only regarding realism and naturalism, but also regarding the theme I shall address here. In a word, just as I would not be wholly identified with Sandro's *Analitico Primo*, in the same way my criticism of the *Straniero* and the *Psicologa* is not to be seen as a criticism to Sandro himself, whose ideas I am not sure coincide altogether with theirs.

2. In the course of his long activity, Quine has given several motivations of his refusal of analyticity. On the whole, we can say that above all he advanced two types of criticisms: (*i*) semantic-pragmatic criticisms aiming to show that it is not possible to explicate analyticity in terms of dispositions to overt verbal behavior; and (*ii*) epistemological criticisms linked to a two-fold consideration: that there are no statements which, like supposed analytic statements, are devoid of empirical content (anti-reductionist and anti-phenomenalistic holism), and that there are no statements which can be considered true regardless of whatever happens. According to Quine, in order to re-establish the accord between experience and the complex of our beliefs, it is possible to revise the truth-value of whatever statement, supposed analytic statements included (revision argument).

With the passing of the years, however, Quine did not acknowledge the same motivational value to these different kinds of criticisms. Here I cannot expound

which two have particular value; on the one hand, the kind of realism 'encapsulated' in the usual conception of knowledge as correspondence which is also at the basis of Tarski's well-known definition of truth; on the other hand the coherentist (or semi-coherentist) character of epistemic justification and then of criteria of truth. Only by neglecting conceptual tensions such as these is it possible to defend answers that are so seemingly straightforward and substantially deficient from one or several points of view. I think answers able to eliminate such tensions can only be given by 'ascending' to the level of conceptual explication. For this reason, criticisms of answers such as mine should not forget their explicative nature. Otherwise, as already noted, they reduce themselves to more or less coarse cases of *begging the question*.

all the various changes in his position.[5] It will suffice to recall the general direction of his changes. At the beginning, the accent was placed on his nominalistic and extensionalistic scruples against the admission of abstract entities such as meanings and his charge against intensional (and then also intentional) entities of being *creatures of darkness* (Quine 1956: 188), devoid of identity criteria specifiable in extensional and behavioristic terms. Later, though, epistemological motivations prevailed. In fact, such a prevalence had already begun appearing in the Fifties. In the course of discussion with Sir Peter F. Strawson, Quine pointed to the reductionist-phenomenal conception of the relation between theory and experience as the main source of plausibility of the analytic/synthetic distinction, for which reason the abandonment of reductionism would also indirectly show the epistemological groundlessness of such a distinction (Quine 1953: 138). Nevertheless, after Strawson's and H.P. Grice's demonstration (1956) that it would be possible to give up reductionism and phenomenalism without giving up synonymy and, through that, analyticity, Quine stressed again the empirically spurious character of intensional and intentional notions as not scientifically reconstructable on the basis of observable verbal behavior.[6]

Such a motivation, though, has progressively lost its strength as two faults of Quine's position have emerged ever more clearly: first, that he made his rejection of analyticity depend upon an, at least very disputable and perhaps unsustainable conception of language, *i.e.* on a behavioristic, naturalistic and ultimately, fundamentally physicalistic conception; second – as Jerrold J. Katz showed – that it was possible to develop an empirical test which linked analyticity to some traits of the speakers' linguistic behavior (see Parrini 1976: I/6). Not for nothing, just when replying to Katz, Quine started again stressing the epistemological motivations of his rejection. In fact, he maintained that on the basis of Katz's operational test, which aims to distinguish between obvious truths of a factual kind and obvious truths based on meanings, "in the really interesting regions – notably in scientific theories – where philosophers have trouble sorting out the analytic sentences, none would count as analytic" (Quine 1967: 53*f.*). In this way, Quine concluded, "Such point as the notion of analyticity was once supposed to have for the philosophy of science would in this way be largely forfeited" (Quine 1967: 54).

[5] What I will say about Quine's criticism of analyticity and intensional (and intentional) notions reflects what I maintained since an essay dated 1973 and republished with some modifications and additions in *Linguaggio e teoria* (Parrini 1976: I). Just on the basis of the ideas expressed in such an essay, today I feel I share the substance of Westphal's (2015) general reconsideration of the analytic tradition.

[6] In Quine (1951: 37), the analytic/synthetic distinction was already described as "an unempirical dogma of empiricists, a metaphysical article of faith."

To me such an answer has always seemed odd (see Parrini 1976: I/6). In fact, from the very beginning some epistemologists (and especially Logical Empiricists) had appealed to the notion of analyticity, or truth on the basis of meanings, just to settle the controversies regarding the epistemological status of some statements belonging to logic, mathematics and empirical sciences. Quine's seminal "Truth by Convention" (1936) seemed to maintain that resorting to such a notion could not give any explicative advantage due to the spurious character of the supposed distinctions traced to, or explained on, its basis. In his reply to Katz, instead, he said that, also in case we succeeded in empirically legitimating the concept of analyticity and the distinctions based on it, such a concept could not be useful to philosophy of science because the epistemological status of the epistemologically interesting principles is uncertain; hence, these principles could not be classified as analytical or synthetic at first sight through an empirical test of Katz's type. However, this fact is one of which we have always been aware. In fact, from the start there have been discussions of the epistemological status of principles such as the axioms of Euclidean geometry, the causal principle or the principles of Newton's mechanics just because it was difficult to classify them. The problem consisted exactly in ascertaining whether it would be possible to clarify this question by recurring to a notion – analyticity – whose application to other kinds of statements did not seem problematic.

Be that as it may, Quine's reply to Katz – however odd it may (not) be – is that the notion of analyticity is devoid of epistemological relevance. Hence, it seems that eventually he saw the deepest and most considerable ground of his criticism of the analytic/synthetic and *a priori/a posteriori* distinctions just in this thesis. This interpretation is confirmed by a reply Quine gave to Geoffrey Hellman in 1986. On this occasion Quine wrote:

> I now perceive that the philosophically important question about analyticity and the linguistic doctrine of logical truth is *not* how to explicate them; it is the question rather of their relevance to epistemology. The second dogma of empiricism, to the effect that each empirically meaningful sentence has an empirical content of its own, was cited in 'Two Dogmas' merely as encouraging false confidence in the notion of analyticity; but now I would say further that the second dogma creates a need for analyticity as a key notion of epistemology, and that the need lapses when we heed Duhem and set the second dogma aside. (Quine 1986: 207)

It appears, then, that eventually Quine came back to his 1950's position according to which the main ground of his rejection of analyticity is the holistic conception of experimental control, in other words his refusal of the dogma of reductionism, even if attenuated. It would be the untenability of reduction-

ism that makes it epistemologically vacuous to speak of statements devoid of empirical content and non-revisable in the light of experience, as the supposed analytical statements should be. In such a way Quine's view of the relation between holism and analyticity could be considered settled, but for the fact that in 1991, on the occasion of the 40th anniversary of "Two Dogmas of Empiricism" and five years after the publication of his reply to Hellman, Quine came back to this topic in a paper titled "Two Dogmas in Retrospect," dedicated to a retrospective appraisal of the theses supported in the (1951) essay.

In this new paper Quine surprisingly seems to limit the validity of holism, in other words: the very reason for which the notion of analyticity should be considered epistemologically pointless. He says that he "regrets" his "needlessly strong statement of holism" according to which "The unit of empirical significance is the whole of science. […] Any statement can be held true come what may, if we make drastic enough adjustments […]. Conversely […] no statement is immune to revision" (*cf.* Quine 1951: 42*f.*). According to Quine, his initial formulation of holism "diverts attention from what is more to the point: the varying degrees of proximity to observation, the example of the brick houses in Elm Street" (Quine 1991: 268). "In later writings" – he adds – "I have invoked not the whole of science but chunks of it, clusters of sentences just inclusive enough to have critical semantic mass. By this I mean a cluster sufficient to imply an observable effect of an observable experiment condition" (Quine 1991: 268).

These statements from 1991 appear clearly to contradict what he affirmed in "Two Dogmas." Now Quine says that not only there are "varying degrees of proximity to observation" – a thesis he had always maintained – but also that it is possible to speak of "clusters of sentences just inclusive enough to have critical semantic mass." This sounds like a proper retraction of the holism defended in the (1951) essay where he *apertis verbis* asserted that "The *unit of empirical significance* is the whole of science" (Quine 1951: 42, emphasis added). That holism seems to disappear if there are groups of statements, no matter how large, which enjoy a certain grade of *semantical* autonomy with regard to the totality of our beliefs. Yet, just when Quine makes this palinode – even, in order to introduce such a palinode – he declares that the holistic pronunciations of "Two Dogmas" from which he is departing are still to be considered "true enough *in a legalistic sort of way*" (Quine 1991: 268).[7] This means that Quine, at least *in a legalistic sort of way*, has not intended in the least to back away from his 1951 assertion that "A conflict with experience at the periphery occasions readjustments in the interior of the field" (Quine 1951: 42) so that, because of the logical and non-logical interconnections among the statements,

[7] [*Ed. note*: a 'palinode' is a retraction.]

there are no parts of the field which *in principle* (in other words, *in a legalistic sort of way*) cannot be involved in the change. Hence, for him it still remains true – at least *in a legalistic sort of way*, *i.e.* in principle – that it is always the whole our "world system" (an expression used by Quine himself in *Philosophy of Logic* [1970: 157], explicitly referring to Newton) that is subject to the test of experience.

3. Although the historical importance of Quine's thought is beyond question, I think that, while reconsidering the epistemological component of "Two Dogmas," instead of unclearly attenuating the holistic conception, he would have better abandoned the formulations which had linked holism to the reductionist version of the verification principle. In fact, such formulations had caused undue over-lappings (not to say confusions) between linguistic and epistemological holism[8] and between language and theory, and hence had obscured the distinction between the linguistic system of a speaker and the complex of his beliefs – as was noted by Noam Chomsky (1969).

What I intend to suggest is that perhaps it would have been opportune to engage in self-criticism about the idea that the two dogmas of empiricism are *at root identical* (*cf.* Quine 1951: 41),[9] and to admit that from an epistemological point of view the acceptance of analyticity can go together both with anti-reductionism, as Grice and Strawson clearly showed in the cited essay, "In Defense of a Dogma" (1956), and the *revision argument*, as Carnap maintained since the years of the *Logische Syntax der Sprache* (*cf.* Parrini 2006: 192-194, 198-203). Carnap had repeatedly noted, in fact, that if 'analytic' means true by language, it is possible to remove the seeming inconsistency between the *revision argument* and the admission of analytical statements by distinguishing be-

[8] I think that the confusion of, or overlapping between, linguistic holism and epistemological holism (which I have already dwelt upon in various passages of Parrini (1976)) is one of the most unsustainable legacies of the so called "linguistic turn" in philosophy, both in general and in its specific Neo-empiristic version characterized by the two interconnected doctrines of the principle of verification and the linguistic theory of the *a priori*. Today, both the linguistic turn and its Neo-empiristic version are largely set aside. Yet even after the end of Logical Empiricism, around the middle of the 1950's, surreptitious forms of such a confusion or overlapping continued manifesting their effects; consider, for instance, Donald Davidson's criticism (1984) of the so called "third dogma" of empiricism (*i.e.* the dogma of the distinction between scheme and content), or the discussion of Quine's holism developed by Michael Dummett and his semantic reformulation of the realism/anti-realism contrast (see Parrini 1998: xv-xvii, 50*ff*). Some Italian effects of Dummett's position are critically examined in Corvi (2010: 189-192).

[9] Regarding this famous Quinean affirmation, allow me to mention a significant episode in Konstanz (1992), during the conference for the centennial of Carnap's and Reichenbach's births. Quine was there; when I quoted the words at issue in order to contest them, he said he had never maintained such a thesis. The astonishment was great, but the audience granted I was right in saying he had.

tween changes which do not involve a change in language and changes which can be classified as changes of the semantic rules of the linguistic apparatus of reference; see Carnap (1963: 899*f.*); Parrini (2006), (2002: chapters 4, 6, 10). What seems impossible to do, instead, is to maintain the holistic thesis in its semantic form linked to the verification principle, according to which the unit of empirical significance is science in its entirety, and at the same time to speak of statements which have such a "critical semantic mass" that they are empirically self-sufficient. If such groups of statements exist, it becomes impossible to continue saying that the unit of empirical significance is the totality of our affirmations about the world!

Additionally, Quine himself once characterized epistemological holism without invoking the verificationist theory of meaning, and hence without appeal to any notion of meaning. It is possible to gather this from a passage of *Word and Object*, *i.e.* a work in which, by fateful irony, the untenable overlapping of epistemology and semantics reached its climax. There Quine states:

> What comes of the association of sentences with sentences is a vast verbal structure which, primarily as a whole, is multifariously linked to non-verbal stimulation. … In an obvious way this structure of interconnected sentences is a single connected fabric including all sciences, and indeed everything we ever say about the world; for the logical truths at least, and no doubt many more commonplace sentences too, are germane to all topics and thus provide connections. However, some middle-sized scrap of theory usually will embody all the connections that are likely to affect our adjudication of a given sentence. (Quine 1960: 12*f.*)

4. As far as I am concerned, to "logical truths" and "commonplace sentences" I would explicitly add both mathematical statements and the statements that we usually deem to be analytic in the sense of being only dependent upon the common linguistic use, however uncertain, vague or richly nuanced it may be. Coming back to our topic, though, I would say that in the case of holism, as in others (in particular, in the case of truth and the option between pragmatism and realism), Quine, for lack of epistemological analysis, has not been able to distinguish two different questions: The appraisal of holism from a logical-epistemological point of view and the appraisal of holism from a practical-operative point of view.

From the logical-epistemological point of view, what counts is the *legalistic* position according to which in principle the whole *web* of our beliefs, due to the interconnections linking those beliefs to one another, faces "the tribunal of sense experience not individually but only as a corporate body" (Quine 1951: 41). From the practical-operative point of view, instead, what counts is

the consideration that what we really aim at in the individual, actual contexts of research is the empirical evaluation of restricted and homogeneous groups of hypotheses, and in many cases of individual hypotheses, often examined in relation to another single hypothesis seen as the only plausible alternative we must keep under control (by observation or experiment). It is important to note that both these dimensions were already implicitly contained in the way in which Duhem presented holism, although he only referred to the *théorie physique*. In fact, Duhem (1906-1914: II/6) criticized the possibility of *experimenta crucis* and their conclusive value not by proposing methods of empirical control which differ from the usual ones, but by pointing out the *logical* impossibility of excluding all the possible explicative hypotheses which are alternative to that which has been accepted, because, *according to the rules of logic,* these hypotheses are potentially infinite in number. Duhem (1906-1914: 329*ff.*) also clearly maintained that there is no logical criterion on the basis of which we can determine which hypotheses are involved in a real, specific experimental test and which hypothesis we must accept or refuse on the basis of the result of that experiment. Over such choices and decisions good sense (*bon sens*), not logic, lords.

If this way of putting things is accepted, it becomes clear that acknowledging the validity *in principle* of holism does not clash with the fact that at a *practical-operative level* we put various sizes of groupings of hypotheses and statements to the test of experience, and in many cases even individual hypotheses or assertions. Holism, in any tenable form, does not require that we disavow this way of behaving, that we declare it to be illegitimate or devoid of any value. Nor does it require that we must search for a mysterious and presumably unreachable alternative procedure involving the system of our beliefs in its totality. An experimental test is always selective, it is always characterized by a certain degree of specificity, determined, I would suggest, by that same good sense to which Duhem appealed for choosing the hypothesis considered confirmed or not confirmed. Therefore, the holistic thesis according to which *in a legalistic sort of way* what is involved is the whole complex of our beliefs is completely valid. This point must be kept in mind as a *memento* that the 'cut' that we explicitly or tacitly make to conduct any empirical-experimental research has an hypothetical value and could turn out to be mistaken. In fact, such a 'cut' is the result of a selection which rests on hypotheses (depending on convictions rooted in good sense and in so-called, scientifically informed "background knowledge") which delimit what *at the moment* we suppose to be important in given specific experimental contexts.

The holistic conception tells us, then, that the conclusion that we have reached is to be considered hypothetical, temporary and revisable. Not only

does such a conclusion rest upon protocol statements which in their turn could be subjected to test and turn out to be unreliable (recall Neurath's metaphor of the sailors), it is also based on a certain way of extrapolating the sub-system of those beliefs which in a specific context we deem to be involved in the empirical control, and so are distinguished from the total system of our beliefs – and it seems impossible to give any absolute foundation of such an extrapolation. In fact, we can never be sure of having rightly selected the sub-system of hypotheses involved in the experiment, nor of having singled out all the genuinely relevant hypotheses. As the history of science has taught us, the most problematic and insidious hypotheses are those which have tacitly operated for a long time, of which we were unaware and which we were not able to include explicitly among those involved in our experimental procedures (recall special relativity theory and Einstein's analysis of simultaneity). On the other hand, it is also impossible to appeal to the fact that in some contexts neither an hypothesis nor its possible competitors can be empirically controlled except by taking for granted one (or more) common presupposition(s), to which we cannot see any alternative. This fact does not confute the holistic thesis of the *logical* impossibility of subjecting an individual hypothesis to control. It only shows that we are not always able to conceive assumptions which are different from those presupposed at the moment – think of Kantian synthetic *a priori* principles and in particular the case of Euclidean geometry before the creation of non-Euclidean geometries (*cf.* Parrini 1976: 192).

Although it is true that Duhem limited his attention to physics, holism in its radical form is only the natural extension of Duhem's idea that when a scientist decrees the falsification of an hypothesis in the light of an experimental result, he can do that only by taking for granted (implicitly or explicitly) the validity of *all* the statements involved in his reasoning or in his argumentation. So it is impossible to deny, although Duhem does not clearly express such a consequence, that *in a legalistic sort of way* among such statements there are also those most general principles, common to all the various disciplinary fields, to which Quine will refer to maintain the validity of holism from a logical point of view, that is from the point of view we cannot leave out of consideration when developing an epistemological discourse; and likewise when one thinks – as do I (Parrini 2018: §5) – that epistemology cannot be deprived of authority by logic.[10]

As I have already suggested, in the case of presuppositional assumptions in the first instance the reference is to those famous Kantian synthetic *a priori*

[10] See Parrini (2018: §5). On this point I refer to Westphal's works (2017, 2018; forthcoming) to underline the importance of "cognitive semantics" for a suitable epistemological theory of non-formal systems of empirical sciences.

principles apodictically certain which are no longer much favored in contemporary epistemology. Some epistemologists, though, have revalued the idea (although in different ways) that the process of epistemic justification (of which empirical control is an integral part) requires admitting such assumptions, but, unlike Kant, in a form which takes into consideration the hypothetical and revisable character of every component of our knowledge. In fact, in order to understand the structure of the relation between our beliefs (scientific beliefs, especially) and experience, it is not enough to distinguish the analytic from the synthetic in an over-simplified way with the analytic intended as including logical-mathematical truths. Not only it is also necessary to consider if and in which measure analytical statements intended as truths in virtue of meaning can include logical and mathematical truths; in addition, and above all, it is necessary to take into consideration that in scientific knowledge, though also in commonsense knowledge, presuppositional principles play a fundamental role, and that the validity of these principles, although not completely independent of experience, depends not on individual, specific experiences, but on experience considered in its totally. These presuppositional principles make empirical knowledge possible, because only through them does it become possible to link individual statements to specific experiences and then to proceed to the usual attempts at empirical control. This aspect of the question, already implicit in Duhem's criticism of Poincaré's conventionalism (characterized by a linguistic inflection),[11] started emerging with Schlick's and Reichenbach's reflection upon the philosophical meaning of relativity theory and the associated doctrine of coördination principles; it presented itself again with Kuhn's so-called 'paradigmatic propositions' and finally resulted in the proposal of a relativized *a priori*, for some scholars linked to Carnap's doctrine of *linguistic frameworks* and for others to the conception of a *synthetic* relativized *a priori*.

5. In my opinion, Quine's epistemological criticisms taken alone (in other words, leaving aside the semantic-pragmatic criticisms) were not such as to invalidate the analytic/synthetic and *a priori/a posteriori* distinctions. Given the structure of epistemic justification, above all in the case of theory/experience relations, they should have led not to the rejection of such distinctions, but to the recovery of a functional, relativized version of them, which version, beside the analytic *a priori* and synthetic *a posteriori*, could also acknowledge

[11] In the conceptual itinerary from Poincaré's conventionalism with its linguistic inflection to Logical Empiricists' semantic-epistemological conceptions, this point – which I cannot examine here – is linked to the reasons that led to the aforementioned, harmful overlapping of semantic holism and epistemological holism. I discuss aspects of this topic in Parrini (1983: 86-90, 96-99, 109-112).

a role for a synthetic relative *a priori*. However, today it is not my intention to speak about this question which I examined in many works.[12] Nor it is my intention to speak about the possibility of inserting the relativized synthetic *a priori* into a model of epistemic justification that can allow us to eliminate the problematic concept of intuition, which Carnap wished as much as Quine (see Parrini 2002: ch. 10, esp. §3). Carrying on my dialogue with Sandro, I would like to show the reasons why it seems to me that the rescue of these and other epistemological notions is worthwhile.

One of the characters of the *Nottola di Minerva*[13] raises just this problem. At a certain point of the book, the *Straniero* discusses the problem of the definition of 'mental', maintaining that, at least for the time being, he would distinguish such a definition from other empirical questions concerning the same notion. This gives the hint to the *Analitico Primo* to ask him ("ironically") whether, by introducing such a dichotomy between definitions and empirical questions, he does not also reintroduce "the distinction between analytical judgements and

[12] My conception of a relativized or contextualized synthetic *a priori* dates back to 1976 (Parrini 1976: 153-290, esp. 264-290). Some decades after, a similar idea has been advanced by Michael Friedman on the basis of an "intellectual" or "historical narrative" (Friedman 2012: 51 *n.*18) centered on the developments that led from Newtonian physics to relativistic theory. As regards Friedman's treatment of this topic, Noretta Koertge (2010, 511*ff.*) has referred to Ernan McMullin to point out that "mechanics is not the only – and perhaps not the best – example to look at when we study the structure of science"; other pertinent critical observations on Friedman's approach can be found in Thomas Mormann (2010). It is also to be noticed that Friedman's conception – that replaced his previous, realistic vision of the philosophy of space and time – initially took the form of a recovery of Carnap's idea of the *linguistic frameworks*. Later Friedman, through a series of "twists and turns" (see Uebel 2012: 7-17, and Friedman 2012: 53 *n.*24), has swung about the way of intending the *relativized a priori* but always remaining linked to a narrative framework of an historical kind tinged with vaguely Hegelian 'necessitarian' connotations. Such connotations are accentuated in Robert DiSalle, who, however, declares himself in favour of a relativized synthetic *a priori*, although he does not give specific reasons for his decision (2010: 524*f.*, 545). As far as I am concerned, from the very beginning I have referred to the position advocated by Reichenbach in the early 1920s and maintained the necessity of admitting a relativized *a priori* of a synthetic kind beside a relativized *a priori* of an analytic kind. In my conception this idea is organized into a vision of scientific rationality that aims at coherence both with what we know about the historical development of philosophical and scientific thought when these are considered in the full variety of their aspects and ramifications, and the firm affirmation of the historically contingent character of every *a priori*, in other words of the contingent character of the cognitive synthesis (*cf.* Parrini 2002: ch. 9); for a general evaluation of this topic I refer back to my recent re-examination of geometrical conventionalism (Parrini 2011: ch. 3). Such a conception is also linked to the way in which I have justified the *a priori/a posteriori* distinction against Quine, and to my position on truth, externalism and realism. On all that, see Parrini in Ferrini (2015), and Westphal in Ferrini (2015: 70-72, 78-79); see also Westphal (2017), (2018) and (forthcoming): in such essays Westphal has very well caught the reasons why I believe that the relativized *a priori* cannot be purely linguistic, as Thomas Uebel maintains (2012: 15-16), nor merely meta-linguistic.

[13] On my use of the theses expressed by the characters of the *Dialogue*, see the *caveat* at the end of §1 above.

synthetic judgements." The *Straniero* replies ("rather wearily"), "No, no! After all definitions too are of course empirical hypotheses. Only after all, though!" Moreover, Sandro makes even clearer the reason that explains the *Straniero*'s weary tone through what the *Psicologa* says immediately after. Unlike the *Analitico Primo*, she intends to face the specific scientific core of the question. The *Psicologa*, in fact, expounds her point of view "with a look that betrays her indifference to … epistemological squabbles," clearly referring to what was said by the two other interlocutors (Nannini 2008: 73).

I think that at this point of his dialogue Sandro touches on an important problem in a two-fold sense, both particular and general. In a particular sense, the *Psicologa*'s attitude gives plastic expression to a thesis of Quine's, *i.e.* the idea that after all in the course of a scientific discussion what really matters is what the scientist affirms or denies, apart from having established at an historical-epistemological level whether we are speaking about definitions or hypotheses. What really matters is only the scientific validity of the result we are trying to establish.[14] Regarding this point, I too am convinced, just like Quine, that in the case of many scientific debates epistemological considerations can turn out to be idle or irrelevant because things are going very well without them. I am also convinced that many philosophical ideas have scant or even nil scientific relevance because they are not properly theoretically or empirically conceived (from this point of view, also philosophers of mind and metaphysicians with an analytical background are not always faultless). I only add, though, that such a position could be shared also by the *Analitico Primo*, if this character is conceived as supporting the old-fashioned Neo-empiristic ideas.

[14] Consider the following passage of *Word and Object*: "Thus it is that in theoretical science, unless as recast by semantics enthusiasts, distinctions between synonymies and 'factual' equivalences are seldom sensed or claimed. Even the identity historically introduced into mechanics by defining 'momentum' as 'mas times velocity' takes its place in the network of connections on a par with the rest; if a physicist subsequently so revises mechanics that momentum fails to be proportional to velocity, the change will probably be seen as a change of theory and not peculiarly of meaning" (Quine 1960: 57). From the point of view of scientific change, Quine's thesis seems to be hardly refutable. Nevertheless, if we take into consideration the logical structure of epistemic justification and in particular the logical structure of empirical control, we clearly see that they involve a functional distinction among the various kinds of the assertions involved and that among such distinctions there are also those between the analytic and the synthetic and between the relativized *a priori* of an analytic kind and the relativized *a priori* of a synthetic kind. In fact, the studies on empirical control have variously shown that the only way of separating the sub-systems provided with a critical empirical (and not semantical) mass from the complex of our beliefs is to acknowledge a functional role to such distinctions, in particular to the distinction between the assertions or principles to which we attribute an indirect empirical content, and those assertions to which we attribute a direct empirical content. Eventually – as we have already seen – Quine himself was obliged to admit that something did not work in his way of conceiving holism, although in my opinion he failed to see the key weak point of his conception.

Coming to the general sense of the problem raised by Sandro, in other words to the fact that scientists often deem, so to say, in principle philosophical and epistemological debates to be fruitless and irrelevant to on their work, I wish to point out at least two reasons for thinking that attitudes such as the *Psicologa*'s are not wholly justified and can turn out to be self-defeating from a cultural point of view.

First of all, history proves that often purposes of purely philosophical clarifications – purposes that were considered of a foundational type when research on foundations was still deemed possible – can turn out to be useful for the development of scientific knowledge. The debates about the nature of space and time or the epistemological status of geometry and mechanics which accompanied the birth of conventionalism and relativistic physics provide the most evident proof of this. Such debates fostered the rise of the theory of relativity and are still useful to better understand both that theory and its relations to other theoretical constructions such as quantum mechanics. This point of view was maintained by Einstein himself, who more than once acknowledged the role of epistemological debates for the birth of the theory of relativity, contributed to those discussions himself and once wrote that "The reciprocal relationship of epistemology and science is of noteworthy kind […] Epistemology without contact with science becomes an empty scheme. Science without epistemology is – insofar as it is thinkable at all – primitive and muddled."[15] In my opinion this means that the coöperation between science and philosophy so dear to the *Straniero* (and Sandro) can easily go together with the work of philosophers chiefly pursuing aims of epistemological clarification.

The second reason highlights a trait typical of philosophy as it has been traditionally conceived and practiced. I think that it would be dangerous to renounce it and limit oneself to the, albeit indispensable, analytical side of the philosophical inquiry.

Science is an essential part of our cultural system and by its purposes and methods it aims at distinguishing itself from other relevant parts of this same system, such as the different forms of artistic expression or religious beliefs. The salient point on which this distinction rests is the high cognitive value (I do not say the exclusive cognitive value, as a dogmatic follower of scientism would claim) that at least *prima facie* is acknowledged to science in comparison with other parts. Nevertheless, we all know very well, whether we are philosophers or not, that such a value, the cognitive value of science,

[15] See, for example, Einstein (1949: 684). Einstein has also made fundamental contributions to the discussion of the geometry/experience relation. Such discussion has such general and autonomous epistemological relevance that still today it cannot be neglected.

is far from being obvious and can be questioned both *in toto* and in part by means of the most various arguments and the most various reasons. In such a discussion – whether one likes it or not – *also* philosophy gives its opinion. Of course philosophy cannot address these questions without adequate knowledge of the object at issue – I mean science – but at the same time, still today, one cannot help referring to the problems traditionally raised about the possibility of knowledge, which are philosophical problems in the proper sense of the term.

Although some epistemological 'squabbles' may appear to be cunning or irrelevant to research scientists in their fields, these same 'squabbles' could have (and sometimes did have) a constitutive function in settling important controversies. Think of the crucial role played by the notion of objectivity in the discussions between relativists and anti-relativists, pragmatists and metaphysical realists, transcendentalists and constructive empiricists.[16] It is difficult to believe that it is possible to take a position on questions concerning the paradigmatically cognitive character of scientific activity and the reliability itself of the results of scientifically conducted researches in comparison to other kinds of statements (for instance, of a religious or mythological type), without referring to epistemological issues concerning the cognitive value of statements, hypotheses and scientific theories, however irrelevant such issues may appear from the point of view of the specific scientific problems addressed on the empiric-experimental level or (also) at the theoretical and logical-mathematical level.

One requirement still integral to today's cultural world, which philosophy first and foremost is called to address, is just an overall vision that could be a reference point for the answers given to problems such as that concerning the reliability and the cognitive value of science. That philosophy must accomplish this task in a way both scientifically informed and conceptually clear and organized seems to me beyond question; that it must renounce this task only so as not to elicit manifest indifference or boredom, not to say nuisance, from scientists (and even some philosophers), seems to me to be only to renounce its very nature.

[16] For example, I think one of the weakest points of the defense of objectivity attempted by Paul Boghossian (2006) in *Fear of Knowledge: Against Relativism and Constructivism* is insufficient attention devoted to the most significant developments in the debate about the logical structure of empirical control to which, after Logical Empiricists, some of the most important post-Popperian and post-Kuhnian epistemologists have contributed.

References

Boghossian, Paul, 2006, *Fear of Knowledge. Against Relativism and Constructivism*, Oxford University Press, Oxford.

Carnap, Rudolf, 1963, "Intellectual Autobiography," and "The Philosopher Replies," in Schilpp, P.A., ed., *The Philosophy of Rudolf Carnap* (Open Court, La Salle ILL): 1-84, 857-1013.

Chomsky, Noam, 1969, "Quine's Empirical Assumptions," in Davidson, Donald, & Hintikka, Jakko, eds., *Words and Objections. Essays on the Work of W. V. Quine*, Reidel, Dordrecht: 55-68.

Corvi, Roberta, 2010, "Dall'olismo epistemologico al pensiero sistemico: un percorso possibile?", in Ulivi, L. Urbani, ed., *Strutture di mondo. Il pensiero sistemico come specchio di una realtà complessa* (il Mulino, Bologna): 175-195.

Davidson, Donald, 1984, "On the Very Idea of a Conceptual Scheme," in *idem*., *Inquiries into Truth and Interpretation*, The Clarendon Press, Oxford: 183-198.

DiSalle, Robert, 2010, "Synthesis, the Synthetic *A Priori*, and the Origins of Modern Space-Time Theory", in Domski, M, & Dickson, M., eds, (2010): 523-551.

Domski, M., & M. Dickson, eds., 2010, *Discourse on a New Method. Reinvigorating the Marriage of History and Philosophy of Science*, Open Court, Chicago & La Salle, ILL.

Duhem, Pierre, 1906-1914, *La Théorie physique. Son objet – sa structure*, Rivière, Paris 1906; 2nd rev. ed., 1914.

Einstein, Albert, 1949, "Autobiographical Notes", and "Remarks on the Essays Appearing in this Collective Volume", in Schilpp, P.A., ed., *Albert Einstein: Philosopher-Scientist*, The Library of Living Philosophers, LaSalle ILL; 2nd ed., 1951: 1-95, 99-688 (*passim*).

Grice, H.P., & Strawson, P.F., 1956, "In Defense of a Dogma," in *Philosophical Review* 65, 2: 141-158.

Ferrini, Cinzia, ed., 2015, *Approaching Contemporary Philosophical Problems Historically: On Idealisms, Realisms, and Pragmatisms*, Proceedings of the Workshop, "Kenneth R. Westphal & Paolo Parrini: A Real Dialogue on an Ideal Topic," in *Esercizi filosofici*, 10, 1: 1-96, <http://www. openstarts.units.it/dspace/handle/10077/11908>.

Friedman, Michael, 2012, "Reconsidering the Dynamics of Reason", in Suárez (2012): 47-53.

Koertge, Noretta, 2010, "How Should We Describe Scientific Change? Or: A Neo-Popperian Reads Friedman", in Domski & Dickson (2010): 511-522.

Lanfredini, Roberta, & Peruzzi, Alberto, eds., 2013, 2016, *A Plea for Balance in Philosophy. Essays in Honour of Paolo Parrini*, vol. 1 (2013), vol. 2: *New Contributions and Replies* (2016), ETS, Pisa.

Mormann, Thomas, 2012, "A Place for Pragmatism in the Dynamics of Reason?", in Suárez (2012): 27-37.

Nannini, Sandro, 2008, *La Nottola di Minerva. Storie e dialoghi fantastici sulla filosofia della mente*, Mimesis, Milano.

Nannini, Sandro, 2015, "Time and Consciousness in Cognitive Naturalism & Replies to Parrini, Di Francesco & Tomasetta, Maleeh & Stephan, Hoerzer, and Roth," in *Rivista internazionale di filosofia e psicologia*, 6, 3: 458-473, 504-513, <http://www.rifp.it/ojs/index.php/rifp/issue/view/18>.

Parrini, Paolo, 1976, *Linguaggio e teoria. Due saggi di analisi filosofica*, Firenze, La Nuova Italia.

Parrini, Paolo, 1983, "L'empirismo logico e il convenzionalismo di Poincaré," in *idem*, *Empirismo logico e convenzionalismo. Saggio di storia della filosofia della scienza*, Franco Angeli, Milano: 45-118.

Parrini, Paolo, 1998, *Knowledge and Reality. An Essay in Positive Philosophy*, The Western Ontario Series in Philosophy of Science; Kluwer Academic Press, Dordrecht; rev. Eng. ed. of *Conoscenza e realtà. Saggio di filosofia positiva*, Laterza, Roma-Bari, 1995.

Parrini, Paolo, 2002, *L'empirismo logico. Aspetti storici e prospettive teoriche*, Carocci, Roma.

Parrini, Paolo, 2006, "Analiticità e olismo epistemologico: alternative praghesi", in Calcaterra, R. ed., *Le ragioni del conoscere e dell'agire. Scritti in onore di Rosaria Egidi*, Franco Angeli, Milano: 190-204.

Parrini, Paolo, 2015, "Time and Consciousness in Cognitive Naturalism", in *Rivista internazionale di filosofia e psicologia*, 6, 3: 474-479, <http://www.rifp.it/ojs/index.php/rifp/issue/view/18>.

Parrini, Paolo, 2018, "'Logical Empiricism: An Austro-Viennese Movement?', or an Unsolved Entanglement Among Semantics, Metaphysics and Epistemology", in *Paradigmi*, 36: 77-93.

Quine, W.V.O., 1951, "Two Dogmas of Empiricism", in *Philosophical Review*, 60; rev. ed. in *idem* (1953): 35-65.

Quine, W.V.O., 1953, *From a Logical Point of View*, Harvard University Press, Cambridge MA, 2nd rev. ed. 1961.

Quine, W.V.O., 1953, "Mr. Strawson on Logical Theory", in *Mind*, 62; rpt. in *idem* (1966): 137-157.

Quine, W.V.O., 1956, "Quantifiers and Propositional Attitudes", in *Journal of Philosophy*, 53; rpt. in: *idem* (1966): 185-196.

Quine, W.V.O., 1960, *Word and Object*, Harvard University Press, Cambridge MA.

Quine, W.V.O., 1966, *The Ways of Paradox and Other Essays*, Harvard University Press, Cambridge MA, 2nd rev. ed. 1976.

Quine, W.V.O., 1967, "On a Suggestion of Katz", in *Journal of Philosophy*, 64, 2: 52-54.

Quine, W.V.O., 1970, *Philosophy of Logic*, Prentice-Hall, Englewood Cliffs NJ.

Quine, W.V.O., 1986, "Reply to Geoffrey Hellman", in Hahn, L.E., & Schilpp. P.A., eds., *The Philosophy of W.V. Quine*, The Library of Living Philosophers, Open Court, La Salle ILL: 206-208.

Quine, W.V.O., 1991, "Two Dogmas in Retrospect", in *Canadian Journal of Philosophy*, 21, 3: 265-274.

Suárez, Mauricio, ed., 2012, "The Dynamics of Reason Reconsidered: A Symposium on Michael Friedman's 'Relativized *A Priori*'", in *Studies in History and Philosophy of Science*, 43, 1: 1-53.

Uebel, Thomas, 2012, "De-synthesizing the Relative *A Priori*," in Suárez (2012): 7-17.

Westphal, Kenneth R., 2015, "Conventionalism and the Impoverishment of the Space of Reasons: Carnap, Quine and Sellars", in *Journal for the History of Analytical Philosophy*, 3, 8: 1-57.

Westphal, Kenneth R., 2016, "Mind, Language & Behaviour: Kant's Critical Cautions *contra* Contemporary Internalism & Naturalism", in Babür, S., ed., *Felsefede Yöntem/Method in Philosophy*, Yeditepe Üniversitesi Yayinevi, Istanbul: 109-149.

Westphal, Kenneth R., 2017, "Empiricism, Pragmatic Realism & the *A Priori* in *Mind and the World Order*", in Sachs, C., & Olen, P., eds., *Pragmatism in Transition*, Palgrave/Macmillan, London: 169-198.

Westphal, Kenneth R., 2018, "Robust Pragmatic Realism in Hegel's Critical Epistemology: Synthetic Necessary Truths", in *idem*, *Grounds of Pragmatic Realism: Hegel's Internal Critique & Transformation of Kant's Critical Philosophy*, Brill, Leiden: §§132-139.

Westphal, Kenneth R., (forthcoming), "Elective Empiricism or Parsimonious Pyrrhonism? Vetting van Fraassen's Voluntarism".[17]

[17] In the interim, *cf.* Id., *Kant's Critical Epistemology* (Routledge, Oxon & New York 2020), §§66-74 (*Ed. note*).

Focus

Ian Hacking
and the Historical Reason of the Sciences

Introduction

Matteo Vagelli, Marica Setaro

Few contemporary philosophers, if any, have had such leading roles in such diverging institutions as the so-called Stanford School of Philosophy of Science (also home to Patrick Suppes and Nancy Cartwright, among others) and the Collège de France (*Chaire de philosophie et histoire des concepts scientifiques*, 2001-2006) – the latter recalling the Chair held by Michel Foucault in the same institution (*Histoire des systèmes de pensée*, 1970-1984). Few others, like Ian Hacking, have successfully undermined the Analytic/Continental divide, by working on the "trading zones" between these two strands, and forged their conceptual instruments by drawing these latter from different sources and applying them to widely diverse philosophical debates, across natural, social and medical sciences: debates ranging from the problem of induction and proofs and deduction in mathematics to the theories of meaning and truth as well as to the controversy between realism and constructivism in natural and social sciences. It would be difficult to find a debate of the main philosophical schools in the last fifty years that Hacking has not tried to assimilate or to contribute to.

Hacking dedicated four books to probability and statistical reasoning, and they are among his most famous works: *The Emergence of Probability* (1975) decisively contributed to introduce probability as a topic into the history and philosophy of science; *The Taming of Chance* (1990) was included by the Modern Library among the 100 most important 20th century non-fiction works, along with the books of a few other philosophers, such as E.G. Moore, Karl Popper, John Rawls and Thomas Kuhn. *Representing and Intervening* (1983) has become a classic of Hacking's production and it focuses on the philosophy of natural and experimental sciences. *Rewriting the Soul* (1995) and *The Mad Travelers* (1998) are Hacking's main contributions to the history and philosophy of psychology and dissociative disorders. He also wrote a great number of papers and essays on widely diverse topics and problems, ranging from ultracold atoms to child abuse and the poverty threshold.

philinq IX, 1-2021, pp. 115-120
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.270

In one of his most recent interviews, Hacking remarked that Ludwig Wittgenstein and Michel Foucault – two of the philosophers who had most inspired his own philosophy – were "classicized" in a short span of time (Hacking 2014). Hacking intended to highlight how easily philosophers' lively and multifaceted research can be transformed into "history" by their interpreters and commentators. Eventually, Hacking will also be historicized by his interpreters, although perhaps to a lesser degree. Indeed, Hacking's work, like that of his great predecessors mentioned above, has been able to be of interest to a wide range of audiences. Hacking's case studies and, even more, the methodological approach on which they are based, have affected not only philosophy but also psychology, sociology and anthropology, among others. Such a variety is due not only to his insatiable curiosity, but even more to the fact that, according to Hacking, the object of philosophy should be found outside the field proper to philosophical traditions. As Georges Canguilhem would argue, "philosophy is a reflection for which all unknown material is good, and we would gladly say, for which all good material must be unknown" (*Le normal et le pathologique*, 1966). Sciences are *par excellence* the raw material of philosophy, because they allow us to study human reason not in the abstract, but at work, that is, through its most refined productions.

Hacking's production is so rich and varied that it would be fruitless to search for a single overarching theme. And still, most of his projects seem to arise from a same concern, which could be well resumed by the following, apparently naïve, question: how can reason have a history? Or better: how can reason have a history and still aim at being objective? In other terms, how can the objective and scientific status of our claims be, if not weakened, at least redefined by the acknowledgement of their historical and therefore provisional nature? Since the 1990s, Hacking has presented his overall philosophical project as in conformity with Kant's aim to make explicit the conditions of the possibility of objectivity. The defining trait of Hacking's inquiries is that they show how these conditions develop historically. And this should not be considered as a simple *addendum* to Kant. It is rather a challenging task to keep the validity of rational claims together with their historically contingent nature. In this respect, it is in Bourdieu's *Pascalian Meditations* that Hacking finds a particularly appropriate way to explicate this fundamental question:

> We have to acknowledge that reason did not fall as a mysterious and forever inexplicable gift, and that it is therefore historical through and through; but we are not forced to conclude, as is often supposed, that it is reducible to history. It is in history, and in history alone, that we must seek the principle of the relative independence of reason from the history of which it is the product (Bourdieu 2008: 25).

This Focus aims to discuss precisely the problems raised by Hacking's version of Bourdieu's "rationalist historicism". In particular, the first two papers by David Hyder and Manolis Simos-Theodore Arabatzis directly approach the question of the *historical* nature of Hacking's philosophy. Hyder's paper frames Hacking's philosophy within what he calls "analytical historical epistemology", of which he provides a conceptual, albeit critical, genealogy. According to Hyder, Hacking would belong to those philosophers and historians of science who, like Nelson Goodman and Thomas Kuhn, embarked on the daunting task of historicizing and naturalizing Kant, opening up to a set of philosophical positions which are, by the same token, empiricist, nominalist and pragmatist. Hyder's insightful reconstruction assigns a central role to Goodman's riddle of induction and his notion of "entrenchment", which are both central references for Hacking who, however, did not articulate the link between such a twofold interest and his *historicized* philosophy of science. When he claims that concepts "have a memory" and that their meaning lies not in their extension but in the trajectory of their past uses, Hacking indirectly confirms Hyder's understanding of analytic historical epistemology as a form of historical inductivism of concepts (Hacking 2004: 8, 37). It should be noted that Hacking never endorsed a self-description as historical epistemologist, and, perhaps more importantly, he explicitly distanced himself from pragmatism (Hacking 2007). He also does not seem to neatly fit into the nominalist box, because there are also realist elements at work in his philosophy. However, the aim of Hyder's paper is to create a large frame in which Anglo-American philosophical debates come in touch with the discussions of the history of science and "Continental" historical epistemology. By showing that Continental historical epistemology preserves normativity, but cannot convincingly account for its ontology, and that the Anglo-American version is provided with an ontology, but cannot convincingly account for its normativity, the dilemma brings to light a tension that has remained too often unnoticed.

In their paper, Simos-Arabatzis argue that Hacking's works instantiate an historical philosophy of science. Their perspective is based on Hacking's reflections on styles of scientific reasoning, a project that constitutes the backbone of his methodology and which, since its first implementation in the 1980s, has catalyzed theoretical discussions and inspired the research of several historians of science. The authors frame Hacking's considerations on scientific styles within the long-standing debate over the "marriage" between history and philosophy of science, animated in particular by Anglo-American scholars such as Roland Giere, Ernan McMullin and Larry Laudan and others, at least since the publication of Kuhn's *Structure of the Scientific Revolutions*. They convincingly show that Hacking's writings constitute an example of in-

ternal combination of history and philosophy of science, in which the historical and philosophical perspectives fully complement each other and are not merely assembled as pre-given building blocks, as in the main "confrontation model" dominating Anglophone HPS. Contrary to Hyder, they see Hacking as eschewing the prevalent naturalizing trend of HPS and rather opting for a hermeneutic approach in which philosophy provides "a coherent and enlightening ordering of the [historical] record" whose aim is to address those philosophical concerns that may be particularly relevant for our present. The authors engage in a detailed discussion of which conditions may ensure scientific stability and criticize Hacking for not being able to reach a middle ground between complete contingency or randomness and inevitabilism in explaining the stability of scientific styles. In their view, the ahistorical, realist and inevitabilist aspects of Hacking's historiography are connected with the metaphysical quadruple relation of truth, language, meaning and belief, which they consider as a trait of the last phase of his project on styles. In this regard, perhaps, Simos-Arabatzis agree with Hyder, who recalled styles of reasoning and Kuhn's paradigms or disciplinary matrixes as inevitable transcendental reminders of differently naturalized frameworks.

The third and fourth papers examine in more depth the topic of stability as a feature that characterises some sciences more than others, despite (or thanks to) their being historical. Both papers link this specific feature to experimentation and to the "self-vindicating" aspects proper to laboratory science, especially referring to its ability to induce the creation of new phenomena. Massimiliano Simons and Matteo Vagelli discuss Hacking's claim to have fostered a "Back-to-Bacon movement" by introducing experiments as a philosophical subject in the 1980s. They show that Hacking's claim was not isolated and that many other philosophers, historians of science and sociologists expressed similar considerations in the same years. The claimed novelty of the philosophy of experiment is usually taken for granted and rarely discussed with a deeper examination of the larger philosophical aims of its upholders. Although in their analysis the authors question the accuracy of this "invention narrative", they do not conclude that Hacking would have therefore not relevantly contributed to the philosophy of experiment nor that the increased interest in experiments in the 1980s should be dismissed as historically uninformed. They rather encourage a reevaluation of the way we assess the history of the philosophy of experiment and Hacking's position in it. The authors devote particular attention to Hacking's realist argument in favor of the existence of experimental entities and show the function that this argument plays with respect to the different debates into which it enters. The "contextualist narrative" which they propose shows, among other things, that in science there is a kind of stability that can

be reached through the accumulation of experimental results and techniques, which, precisely like the styles of reasoning, cut across different theoretical frameworks. This kind of stability, which is not achieved "in spite of" but rather thanks to historicity, is more apt than the idea of linear progress and that of revolutionary breaks to account for the relative steadiness of natural sciences, at least since the last revolutions of the first half of the 20th century. Hacking finds reasons to believe that Gaston Bachelard's philosophy of experiment and his notion of *phénoménotechnique* at least partially inspired this idea, thus showing that experimental themes run across historical epistemology, broadly understood.[1]

In her paper, Jacqueline Sullivan applies Hacking's ideas on stability to the cognitive sciences, which Hacking himself has relatively neglected, except for rare references, in which he points out that cognitive sciences are mainly sciences that "represent" and do not "intervene". On the contrary, Sullivan shows the plethora of experimental activities through which cognitive neuroscience attempts to draw from rodents' behavior useful insights into human cognitive functions as well as into neuropsychiatric and neurodegenerative impairments. Sullivan's paper provides further evidence to Hacking's claim, also discussed by Simons and Vagelli, that "experiments have a life of their own" and that stability can be reached rather at a level of intervention than at a level of theory and auxiliary hypotheses, as occurred in Duhem's coherentist thesis. Sullivan provides considerable evidence to support the idea that cognitive sciences can reach stability not "despite" but precisely because of their disunited and dispersed character. Such a stability can be reached locally through a "mutual maturing of types of apparatus, phenomena and theories", despite the general lack of conceptual, methodological and explanatory unity that still characterizes cognitive sciences. The study of the same cognitive function can benefit from the diversity of epistemic standards (including background assumptions, methods, vocabularies and materials) applied to it. It is argued that together with the idea of intertheoretical reduction we should also abandon the idea of unity as the search of a single system of scientific classification of natural kinds. "Cognitive kinds", i.e. the cognitive functions under experimental investigation, are phenomena which are created in laboratory settings and whose existence is as fleeting as that of the electron in the cloud chamber. Sullivan does not question the fact that they are real, but interestingly shows that for them, too, stability is given by disunity, that is, by letting fundamentally different practices and methods successfully develop.

[1] See Bachelard's "The dialogical philosophy", presented for the first time in English translation in the Past Present section of this issue of *Philosophical Inquiries*, pp. 231-240.

Albeit well-known and widely inspiring, Hacking is still rarely studied, and his wide-ranging production has not yet received an accurate and comprehensive analysis. This Focus aims to precisely fill this gap, by providing one of the first extensive studies dedicated to Hacking's philosophy. It does not wish, however, to cover all the philosophical areas to which he has possibly contributed, neither does it aim, more generally, to provide a commentary nor an exegesis of his works. By collecting papers by both established and young scholars, this Focus rather intends to explore why Hacking has so largely influenced the field of history and philosophy of science. Analysing Hacking's contribution to 20th century attempts to bring together history and philosophy of science as well as discussing his arguments on scientific stability, the Focus tackles, from different perspectives, the question of the historicity of reason. Without aspiring to definitive answers, this Focus wishes to open up lines of further research *on* Hacking's works as well as *along* their path.

Matteo Vagelli and Marica Setaro

References

Bourdieu, Pierre, 2008, *Pascalian Meditations*, in Hacking, Ian, ed., *Scientific Reason*, Taipei, National Taiwan University Press.

Hacking, Ian, 2004, Historical Ontology, Harvard University Press, Cambridge MA.

Hacking, Ian, 2007, "On Not Being a Pragmatist: Eight Reasons and a Cause", in Cheryl Misak, ed., *New Pragmatists*, Oxford University Press, Oxford: 32-49.

Hacking, Ian, 2014, "Ian Hacking, The Philosopher of the Present. An Interview by Matteo Vagelli", in *Iride*, 27, 72: 239-269.

Naturalism, Pragmatism and Historical Epistemology[1]

David Hyder

Abstract: Historical Epistemology is a discipline that draws on quite distinct sources, straddling the analytic-continental divide within the history and philosophy of science. In this paper, I examine the analytic side of the equation, namely the tradition of empiricist naturalism, and the emergence, within the work of Goodman, Kuhn and Hacking, of naturalized transcendental structures resembling Wittgensteinian language-games, and the correlated multiplication of "worlds".

Keywords: Goodman; Hacking; Kuhn; naturalism; pragmatism; history and philosophy of science (HPS)

Historical Epistemology is a discipline essentially associated with the Max Planck Institute for the History of Science, having been developed as its *raison d'être* by Lorenz Krüger, the driving force behind the Institute's creation, in collaboration with Lorraine Daston and Ian Hacking. Krüger died in the year before the Institute opened, meaning that his legacy was assumed and developed by the first directors of the Institute, including Jürgen Renn and Hans-Jörg Rheinberger. The work of these authors, along with that of dozens of associated researchers and post-docs in the intervening years, has come to define the field, though often without a unified understanding of what the term should mean.

This is certainly no major shortcoming – there is no movement or school of thought that does not, on closer inspection, reveal considerable diversity among the approaches of its members. There is no reason to think that the lack of a consensus among the members of the Vienna Circle on a host of issues somehow indicates a methodological shortcoming, nor that the term "Logical Empiricism" thereby becomes superfluous or misleading. In both cases, the

[1] I thank Alexandre Declos, Matteo Vagelli and the editors for comments and for criticisms of earlier drafts. I also thank the organizers and participants at the « 4ème Journées d'étude: Épistémologie historique et les désunités des sciences » in Paris, May 2018, for their incisive remarks on the lecture which is its source.

ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.342

proper way to respond to this diversity is to examine and compare these approaches and thereby develop a "cluster-concept" that covers adequately both the commonalities and the differences. Such an analysis presents special difficulties in the case of HE, because of the deliberately diverse sources on which its creators drew. Loosely speaking, the tradition has both "analytic" and "continental" origins, for instance Kuhn versus Foucault, and from the point of view of late 20th c. HPS, these authors were quite remote both in their methods and their objects of study. One positive consequence of HE has been a dramatically improved understanding of the pre-war context, and of authors such as Helene Metzger, Ludwig Fleck, Georges Canguilhem, and George Sarton, none of whom fits neatly on just one side of the cleavage. Extending the history of the discipline into the first half of the 20th century allows one to bridge this gap, without forcing any single author into one framework.

In this paper, I propose the second part of what might be called an historical epistemology of "Historical Epistemology". In the first part (2003), I have discussed some of the French and German antecedents to the programme, focussing on the normative transcendental residues we find in authors such as Husserl, Cavaillès, Canguilhem and Foucault. In the following, I examine the other side of the equation, namely the tradition of naturalist empiricism, and the emergence, within the work of Goodman, Kuhn and Hacking, of naturalized transcendental structures resembling Wittgensteinian language-games, and the correlated multiplication of "worlds". It can therefore be regarded as an historical epistemology of *analytic* HE, and it is executed following the method of that discipline itself, as I came to understand this while studying under both Krüger and Hacking in the early 1990s, and while working at the MPIWG in the Department of Hans-Jörg Rheinberger from 1997-2000.[2] Most of the views I ascribe to the above-mentioned authors I have learned less by reading their works, and more in conversation and debate, meaning in turn that their views may have changed, or, in some cases, were never published in the first place.

This accounts for what will strike many readers as an incongruity, namely the central position of the work of Nelson Goodman throughout the following discussion, even though there has been little discussion of his role in the HE literature.[3] This is, I believe, an important omission, since Goodman's *Fact,*

[2] This method is being employed in this paper primarily to *illustrate* what I believe Historical Epistemology *should* look like, if it is done at all.

[3] The most notable exception is Hacking himself. The importance of Goodman for my own work derives from an intensive study of his *Fact, Fiction and Forecast*, which I began in 1993 under Hacking's direction in Toronto and pursued from 1993-1995 under Krüger at Göttingen. In the final stages of drafting this paper, I was made aware of Hacking's (1993), which provides a quite different view of many of the topics explored here.

Fiction and Forecast (*FFF*) not only drove the development of projects such as Kuhn's, but equally well made the topic of counterfactual conditionals a central concern of both philosophers and scientists. Both the "analytic constructivism" we find in Kuhn, and modern theories of counterfactual conditionals up to and including the work of Kripke and Lewis find their source in this book. But the argument presented is of such logical concision that few writers today appreciate the impact that it had.

Similar remarks apply to my discussion of pragmatism and its antecedents. Goodman's own solution to his "New Riddle of Induction" was essentially pragmatic, and the same can be said of the late Wittgenstein, of Kuhn, who drew on both sources, and, finally, of Hacking himself. The *nominalism* that creates the "New Riddle" and the *pragmatism* that solves it are the philosophical basis of the programme of HE in its analytic guise. But, as I will argue by examining the history of these doctrines, they come with certain costs. Most obviously, nominalists are, by definition, anti-realists, meaning that they have programmatic objections to strong ("metaphysical") theories of causality such as emerged late in the 20th century. And if that nominalism is applied just as well to our mental and linguistic representations, as it is in the works of both Hume and Goodman,[4] the result is a near total collapse of normativity. Since both outcomes are desired by the nominalist, we should hardly expect that he will acknowledge either as a shortcoming. He will say that the ultimate norm is always utility and that the ultimate argument for accepting his position must also be pragmatic – what else could it be? But, as I will suggest in conclusion, the empirical evidence speaks against the utility of pragmatism. If one reacts by saying that it is a *normative proposal*, it thereby acquires the same epistemological status as the aprioristic and metaphysical doctrines it seeks to eliminate.

1. *Analytic constructivism: "we live in many worlds"*

Phrases like these are often heard around the HPS community, along with kindred constructivist claims along the following lines: "object X did not exist before Y", "natural kind W was constituted by experimental practice Z", and so on. The claims are deliberately provocative. When pressed, the speaker

[4] Both authors argue that general representations ("abstract ideas") are particulars attached to a word and projected by means of a similarity-relation onto their extensions (the "extension" of a predicate is the set of all objects to which it applies). One way of distinguishing the Humean tradition from the Kantian one is to examine the status of this relation: for transcendental philosophers, it is "internal", in Moore's and Wittgenstein's sense of the term, whereas in the empiricist tradition, it is external.

usually concedes that they are not referring to the Multiverse, or to other galaxies, when they speak of many worlds, but to something else. Or they will say that it was not the *existence* of, say, protons, that was called into question, but the *concept*. Yet to be told that we live in many worlds, but not literally so, or that scientists invent new *concepts* is hardly provocative – the first is merely a metaphor, while inventing new concepts is exactly what theoretical science is supposed to do, and indeed normally does.

Within the Anglophone HPS community, such claims all go back to the work of Kuhn, who introduced the language of paradigm-shifts and semantic revolutions – changes in the conceptual scheme of a scientific community that are so deep-reaching that they may be compared to the acquisition of a new language. Kuhn's defence of this claim, developed in his *Structure of Scientific Revolutions*, bears some resemblance to Foucault's *Archaeology of Knowledge*. Both are works written by historians of science in which they retroactively identify a methodology in their earlier work; and both authors do so by modifying the view of a neo-Kantian predecessor – in Foucault's case, Husserl, and for Kuhn, Carnap. Whereas Foucault retained, if uncomfortably, the normative-transcendental component of this Kantian background, in Kuhn's work the constitutive framework is *naturalized*.

2. *Kuhn*

Kuhn is known well outside his field for two interrelated claims: mature scientific disciplines are governed by "paradigms" and "disciplinary matrices" that determine the methods and problems historical actors can consider; but these disciplines undergo punctuated evolution, marked by revolutionary "paradigm-shifts". *Structure* implements a dynamic version of Carnap's late conventionalism as empirical history, which, once one recalls the neo-Kantian background, leads naturally to the language of multiple worlds. For in that tradition, philosophy is concerned with *foundations* for the natural sciences, that is to say their logico-mathematical and inductive *principles*. As in Kant, these are first "constitutive" (Early Logicism), later they become "conventional" (Reichenbach and, earlier, Poincaré). If and when we change these principles, we change our "framework" or "scaffolding", and since the framework, for any neo-Kantian, plays the same foundational role as traditional metaphysics, it follows that changing the framework "changes the world".

In consequence, Kuhn claims that scientists living in different eras "live in different worlds". The worlds are *defined* by their fundamental normative commitments, which simultaneously *separate* them. These worlds are successive – in each historical phase of a science, terminology and methods

are shared, and so is ontology. After a revolution, however, there is a loss of translatability, which parallels exactly the same failure that – supposedly – accompanies a radical change in a category-system. We are, with respect to our distant scientific forebears, in the same position as a field-linguist encountering an historically isolated dialect within her own language-group.

Seen from this point of view, it is evident that Kuhn's theory is *historico-epistemological* in much the same sense as other authors in the French and German traditions, specifically Husserl's *Crisis of European Sciences*, Canguilhem's *Normal and the Pathological* and the earlier works of Foucault up to an including *The Order of Things* and *The Archaeology of Knowledge*. All of these authors appeal to *historical a prioris*, *epistémès*, *disciplinary matrices*, and the *ruptures* and *revolutions* that separate successive conceptual systems. Since, however, we are focused in this paper on the analytic side, I will turn immediately to the *difference* between the French and German approaches on the one hand, and the American one on the other. This difference derives from a commitment to "naturalist empiricism" which, as we shall see, could be more aptly called *nominalist inductivism*.

Nominalist inductivism is an ancient philosophical position with an equally long history, since it is canonically addressed and rejected in favour of what we now call "Platonism", "essentialism", or "natural-kind realism" in Aristotle's works on logic, specifically *Posterior Analytics* I.i (71a30f.). It is the philosophical thesis that all universal propositions are provisional, because (i) it is impossible to know the extensions of their predicates in advance, meaning that (ii) piecemeal induction on actual past observations is the only source of *credibility* for universally quantified propositions, and such credibility never amounts to *certainty*. At various points in history, (i), (ii), and their ancillae have been called "nominalism", "extensionalism", "empiricism", "naturalism", "pragmatism" etc. Thus the 20th c. version of this position, typified by the work of Quine and those around him, has classical, medieval, early modern and 19th c. antecedents. But it is still more extreme. Almost all earlier authors, even Hume, made exceptions for at least some of the propositions of mathematics and logic. Twentieth-century American nominalist inductivists went further, however. Not only was the entire project of synthetic a priori foundationalism to fall. Even *analytic a priori* principles, including constitutive conventions in the style of Carnap were to be rejected. But Quine's project failed, and the way in which it failed is, one might say, the condition of the possibility of 20th c. analytic historical epistemology.

3. *The condition of the possibility of analytic historical epistemology*

As we just saw, Kuhn's worlds are *successive.* But when we look at much contemporary Anglo-American HPS, we get the claim that we live in multiple worlds *at the same time* – the many worlds are now *simultaneous.* So, what changed? The key figure is Nelson Goodman, the author of *Ways of Worldmaking*, and a collaborator of Quine's in their earliest days. This collaboration began as the project just outlined: by nominalizing the new logics, one would dismantle the last *a priori* remnants in Russell and Whitehead's *Principia*, Wittgenstein's *Tractatus*, Carnap's *Aufbau* and his *Meaning and Necessity*, thereby destroying the analytic foundations of mathematics that were the principal goal of logicism. This forcibly undermined philosophy's *normative* status: since no universal propositions are certain, there are no rational foundations for logic, mathematics, or the physical sciences; there are also no certain universal principles in ethics, including universal human rights. The difficulties Goodman later highlighted in his *Fact, Fiction and Forecast* (*FFF*) emerged only when this radical nominalism was applied in empirical science, where the target was the types of necessary connection involved in *natural kinds* and *causation*.

Quine never abandoned his extensionalist project, but Goodman did quite early on, and he presented his reasons to Quine as a dilemma: abandon (i) set-theoretical extensionalism, or abandon (ii) induction. More precisely, he argued that there was no way of implementing a theory of inductive confirmation without appealing to one of three alternatives: Natural Kinds, Natural Laws, or Counterfactual Conditionals. A solid theoretical account of any one of these, Goodman argued, could serve as a basis for the other two; however, all of them were problem-ridden, because they all involved the concept of *possibility*.[5] Goodman offered his own, fourth solution, that would accord with Quine's and his scruples, because it appeals only to past regularities in the use of names – his theory of predicate-entrenchment. It follows that human and social factors are unavoidably involved in the practice of science at the level of induction itself. It is this last solution that motivates the Kuhnian and post-Kuhnian tradition we are discussing. But it is not the solution most philosophers opted for, at least among the younger generation, and this fact alone accounts for much of the gap between what historians of science today consider to be "philosophy", and what is actually going on in 21st c. philosophy departments.

[5] Thus it is worth reemphasising that any "solution" to Goodman's paradox that appeals to one of these factors is, from Goodman's point of view, a concession that he is right.

4. *Why does nominalism matter to constructivist HPS?*

It is not hard to see why some form of nominalism is essential to those who believe in "made up worlds" or "made up people". Nominalism is the position that denies universal forms and asserts instead that all categories are human creations. A fortiori, it justifies the sort of *ontological pluralism* we have been discussing: if the world is a box of sand, consisting of indefinitely many x's, none of which is possessed of real properties, then all ways of classifying the grains have the same ontological standing. Since nothing has real, let alone essential properties, and all predicates are derived from human thought or language, it follows that the fundamental metaphysical categories, and the constitutive principles of the sciences that mention them, are either inductive generalizations or conventions. Thus, not only is a *succession* of incompatible worlds possible, but such worlds are *simultaneously compossible.*

In the early work of Quine and Goodman, nominalism is a means of deconstructing logical and mathematical foundationalism, which they attempted by completing the "no-classes" theory of *Principia Mathematica*. Russell and Whitehead had argued that, in order to block the set-theoretical paradoxes, we first had to reduce set-theory to logic, and then, second, introduce a hierarchy of functions and logical forms, thereby blocking the formation of the *pseudo-statements or -judgments* that apparently asserted the existence of these absurd sets. But Russell's theory of propositional judgment contained "forms", which were the inevitable residue of the classes he had eliminated, and which reappeared in the definite descriptions that Russell had used to handle non-denoting signs, such as propositions, class-terms and proper-names. So a principal aim of Quine's and Goodman's (1947) "Steps towards a Constructive Nominalism" was to eliminate this residue, which they achieved by means of linguistic behaviourism. This was a natural fit, since the latter doctrine replaces cognitive meaning-bearers (concepts, ideas, intentions) with *dispositions*, which are in turn interpreted as empirical hypotheticals. Concretely, linguistic behaviourism allows one to replace "mental forms" with sequences of tokenings. Instead of presuming the existence of a sensible property such as *red*, which, when wedded to the word "red", determines its extension, we introduce a definition such as: "Set of objects in whose presence English-speaker Q utters 'red'". Property-ascriptions are now language-relative, and, more importantly, they are essentially unbounded: no fact in the world determines in advance the extension of "red", and any claim to know that extension is a *prediction*, that is to say a "projection" of that term onto future linguistic behaviour. The effect is to turn every universally quantified universal proposition into an empirical prediction, including, as intended, the

"logical propositions" at the heart of Frege's *Foundations*, Russell and Whitehead's *Principia*, and Wittgenstein's *Tractatus*.

The premises for this conclusion were: first, that extensions are determined by tokenings, and only past tokenings are known, meaning that there are no future extensions beyond those anticipated by our dispositional "projections"; second, that universal propositions acquire whatever credibility they possess only by induction on past observations, which means, for both Hume and the behaviourist, through the formation of *habits* or *dispositions*. Goodman's paradox then implements the following dilemma: accept a principled distinction between coextensive terms or abandon induction. He does this by arguing that, for any inductive context in which we may find ourselves, there is a projection-ambiguity that cannot be resolved in extensionalist terms. Concretely, we *must* choose between "green" and "grue", and, furthermore, we *do* choose "green". But everything that differentiates them lies in the future or has never been observed. Therefore, our inductive practice is based on convictions that are essentially non-extensional – the nominalist inductivist must choose between the two.[6]

5. *Goodman's rejection of extensionalism*

Given the enormous literature on Goodman's paradox, we will dispense with a statement of it,[7] and will try to avoid entering into disputes regarding its correct formulation. What interests us is the dilemma just outlined, and it does not depend on these finer points, for Goodman's target is very specific. His argument shows that Quine's and his youthful nominalism,[8] however successful it might be

[6] The simplest way to drive home the point is to consider a world like Goodman's which, for whatever reason, is destroyed immediately after time *t*. In this world, the actual (or "manifest") extensions of "green emerald" and "grue emerald" are strictly identical, therefore any property, including causal ones, that is ascribed to the one must be ascribed to the other. That "they will differ after *t*" is a belief that speakers in this world may well hold (prior to *t*); however, nothing "corresponds to" or "verifies" this psychological state (similar remarks apply to unobserved emeralds prior to *t* – the future and the unobserved past are symmetrical with respect to the logical point). For this very reason, it can do no work in parrying Goodman's paradox in the world he considers.

[7] On Goodman's original definition, an object is *grue* if-and-only-if it was examined before some time *t*, and was observed to be *green*, *or* it is *blue*. *Grue* and similar predicates are often called "disjunctive", because they fuse, by means of the *or*-connective, or "disjunction", two apparently "natural" classes of objects such as *green* and *blue*, to generate an absurd class. The New Riddle of Induction challenges us to explain why, at time *t*, we prefer to induce on "green emerald" rather than "grue emerald", given that their actually observed extensions are, by design, identical prior to *t*, but diverge thereafter, and thereby generate incompatible predictions.

[8] As indicated earlier, Goodman's allegiance to this position was always limited. Both of them soon abandoned the extreme position of their (1947), Quine adopting classes, but no natural kinds,

in undermining the claims of *mathematics* to a priori certainty, cannot be reconciled with the needs of *empirical science*. Since the naturalistic conclusion of the nominalist argument was to be that *all* science is fallible empirical science, the naturalist must now pick either (i) his nominalism or (ii) his inductivism.

This point can be brought out by considering our natural responses to Goodman's problem: we could appeal to the fact that "green" refers to a real, as opposed to an artificial kind; when pressed, we could appeal to counterfactuals in order to demonstrate that "grue" is in some sense absurd; finally, if asked how we know that these counterfactuals are true, we could invoke the laws of chemistry and optics, which can be distinguished from "mere regularities" by appealing to a difference between natural and artificial kinds. These options resemble – from Goodman's point of view – a closed curve in space. For each projects beyond an *actual* extension (of *past* tokenings) to a *possible* one. When we say – in the simplest case – that the concept "green emerald" *differs* from "grue emerald", we anchor the difference in something that either *has not yet* happened (an observation of a green emerald after *t*), or that *could have* happened but didn't (a counterfactual observation of a green emerald before *t*). But that something is not part of the actual, or "manifest", extensions of "green emerald" and "grue emerald". Indeed, "grue" has been defined in such a way that these sets are *necessarily* identical, so that no fact of the matter could ever distinguish them before time *t*. Any appeal to supplementary criteria, for instance laws of nature, will also involve general propositions whose terms have open extensions, and the Goodmanian will simply disjunctivize those as well.

6. *Three possible solutions*

The addressee of Goodman's argument is therefore forced to choose between extensionalism and inductivism. Quine stuck to the former and never did, to my knowledge, offer a substantive response to Goodman. Most philosophers and historians, however, made the opposite choice, and the specific *intensionalist* solutions they chose define the field of methods arising within analytic HPS in the second half of the century. The options were: metaphysical realism, counterfactual definiteness, and natural laws, all of which involve some form of realism about universals, and the normative force these have carried since Plato. I will first say a few words about each.[9]

Goodman favouring qualia much like our "tropes" or Wittgensteinian "objects", while rejecting classes (for a discussion of this period, see Cohnitz & Rossberg 2006: 86ff.). It is easy to see how the grue-paradox works in favour of Goodman's choice, and against Quine's.

[9] These options are sequentially examined in the first chapter of *FFF*, "The Problem of Counterfactual Conditionals".

(1) Natural Kinds ("metaphysical realism")

This solution responds to Goodman's paradox by pointing out that green objects form a natural kind, whereas grue objects do not. This is the simplest response provoked by *FFF*, which Goodman no doubt intended to elicit in order to drive home his fundamental point. For not even the scientific realist believes that colours are real, in the sense that they might hold a spatio-temporal distance to be. The reality of colours long yielded its place to such scientific properties, but Goodman's argument applies just as well to them: if the meaning of a predicate simply *is* the set of its instances, then there exist, for every such predicate and at each time *t*, many disjunctive correlates[10] which will recreate the original dilemma. Furthermore, since every appeal to future or counterfactual observations is a projection, there are no *causal* properties that can be ascribed to the "real" properties which do not automatically apply to their coextensive doubles.

(2) Counterfactuals

Counterfactuals seem to have been an early candidate for solving the problems that led to *FFF*, since it begins with a lengthy analysis of what were, in fact, Goodman's own earlier failures to make them work to his satisfaction. But if we had a theory of counterfactuals – for instance that of David Lewis – it might not be difficult to decide the question in favour of *green*. As many authors have pointed out, given an entirely plausible claim concerning some green object G that I did *not examine* before *t*, such as "Had I examined G, it would have been found to be green," I can infer with certainty[11] that, "Had I examined G, it would have been found to be grue." And if objects don't change their properties for no reason, then it seems reasonable to assert that, since it *would have been found to be* grue, it *is still* grue. But that entails that, *were* I to examine it *now*, after *t*, it would be found to be blue, contradicting our original assumption that it is green. Thus either (i) objects change their properties for no reason, (ii) the predicate *grue* is absurd, or (iii) some of the relevant counterfactuals have no truth-values. If we hold to (i) on grounds of simplicity and economy, then either (ii) *grue* is an absurd predicate, or (iii), the negation of Counterfactual Definiteness[12] (CD), must hold.

As already suggested, Goodman appears to have rejected this approach

[10] See footnote 7.

[11] Because this inference follows from the definition of "grue".

[12] That is, we must reject the supposition that statements such as: "If an observation of E had been made, result G would have been obtained" always have a truth-value. See Skyrms 1982: 43 and Stapp 1971 for discussion.

even before formulating the New Riddle, because it involved either possibilia, or an appeal to Natural Kinds. Rejecting the first, Natural Kinds become the obvious option, but then we recur to (1). And, in fact, things have grown still more problematic for this option since Goodman's work, for CD is a variant of Einstein, Podolsky and Rosen's "Criterion of Reality" (1935: 777), and it is only by removing this same realist premise that one can neutralize their argument. In consequence, anyone who goes this route[13] will eventually have to confront the role of counterfactuals in EPR.

(3) Natural Laws

The notion of a Natural Law is stronger than that of an empirical generalization. As Hume complained, the notion of a Law adds something to the latter that goes beyond the prior instances observed, and indeed beyond the future instances as well. The difference corresponds exactly to the Medieval distinction between "sempiternal" (always true) and "essential" (always true by necessity). If we are pushed back, by recursive application of arguments in the style of (1), to our fundamental categories and laws, and if we are able to reject their disjunctivization, then it is an easy matter to show that *green* is legitimate and *grue* is not. But to do so is to claim that one has some way of knowing, once and for all, that one has arrived at these fundamental laws, and the property-concepts they deploy. And on this point, Goodman is intransigent: any appeal to metaphysical properties that *causally necessitate* observable regularities is unscientific (*FFF*, 20). What of, one may therefore ask, the disposition that forms in us when we resolve to project in one direction in preference to another?

7. *Constructivism (back to HE)*

Thus we are left with Goodman's own, fourth option, which is the philosophical foundation of the analytic constructivism which followed. On this approach, I reject realism, counterfactual definiteness and the notion of natural laws, and restrict myself to what has actually been observed. The problem is that the latter does not, by design, offer any facts that could break the symmetry.

Yet all is not lost, Goodman argues, because we have overlooked the fact that the *use of names* falls within what has actually been observed. And, quite clearly, this leads to a form of *intensionalism*, in that it allows us to ascribe

[13] Including this author. For background, see Dickson 2002: 657. Note that there are many versions of EPR in the literature that avoid this premise; however, most involve, for reasons that are too complex to discuss here, Absolute Time.

properties to "green" and "grue" that distinguish them,[14] given that, once again, grue has been so defined that no observation of its *referents* could ever yield such a criterion. But if I am allowed to appeal not only to what my words refer to (extensions), but also to the words themselves, a simple response to the dilemma presents itself: one of the words has been used a lot, the other hasn't. This is a fact about names in their past application to things, and not a fact about the things themselves. It is, therefore, a *nominalist,* but *non-extensional* solution.[15]

Suppose now that I am a nominalist who wishes to explain the difference between green and grue, without appealing to real properties or counterfactuals. I find myself obliged to do this, because there is no science without induction and, given that I have sacrificed the reality of extensions in order to undermine necessary truth, I must offer a substitute. This substitute turns out to be the *past usage of scientific terms.* For, since everything that is not a particular referent (=X), is a human creation, and these "intensional" factors are necessary to science, I also believe that natural-kind concepts of human creation are ineliminable. There is no empiricism free of a human-created foundation, and that foundation is (1) neither a free choice, nor is it, (2) derived from the *nature of things,* for (1a) it precedes me, and (2a) I remain a nominalist. Thus, I must conclude that scientific practice is essentially dependent on an historically contingent conceptual frame – a paradigm, a disciplinary matrix, a style. Put otherwise, Goodman's argument leads unavoidably to the view that concepts have memories.

8. *The Normative vs. the Descriptive*

The argument I have just given is the philosophically interesting and important basis of HE – the one that convinced its philosophical creators. But it's not, from what I can tell, what people understand by this term today. What we get much of the time is *just history*, without a genuine *epistemological* component. This is due to a contingent accident: Lorenz Krüger died just before the Max Planck Institute he worked so hard to establish was founded. Historical Epistemology became the province of Anglo-American historians *only*, many of whom have contempt for philosophy, or are, at best, "naturalized epistemologists" à la Quine. That's the sociological situation. But the problem is system-

[14] Trivially, *green* is longer than *grue*, but that is a property of the *names*, and not of what they refer to. Goodman did not regard this as "intensionalism", but it does agree to our current understanding of, for instance, an "intensional" context. Those who object to the term can substitute "non-extensional".

[15] See Declos 2019 for a discussion of the nominalism involved here.

atic and theoretical, because to *solve* Goodman's problem, we need a definition of "good" concepts (green over grue). And "good" is a normative term.[16] If you try to define it in terms an empiricist would accept, you get *pragmatism*.

9. *Pragmatism: a potted history*

In its modern form, pragmatism is the view that whatever maximizes utility is better than whatever does not. Given a market that accurately prices utility, pragmatism in epistemology yields the view that *what sells is true*. But most contemporary pragmatists resist this natural connection to economics, preferring to leave open the nature of the interests in question, and the measure of utility. Thus, when philosophers in this tradition tell us that "It's real if you can spray it," they mean "if you can *do* [Gr. *prattein*] something with it", where the nature of this "doing" is at first left dangling. But whatever it is, it is going to be something in the human life-world, something connected to human interests and desires, without which the notion of utility in question draws a blank. It might seem at first glance as if pragmatism in the philosophy of science is a special case, since today we have pragmatist theories of ethics, epistemology, etc. But if we engage in a little history, we discover that it emerged hand in hand with the Early Modern version of nominalist inductivism that Kant had named "Empiricism", specifically within the collaboration between John Locke and Thomas Sydenham on the theory and practice of medicine. Already in these early authors, pragmatism plays the role of metaphysics. For the empiricist, having rejected appeals to natural kinds and hidden mechanisms, owes us an explanation of our realist convictions, and this explanation must give account of their epistemological role.

Kant's choice of the term was not friendly, and the view he ascribed to "empiricists" was closer to that of Berkeley, and indeed to 20th c. sense-data verificationism, than it was to Locke's, or even Hume's views.[17] Locke did not deny the existence of a mind-independent world, nor even our ability to learn things about the Cartesian sea of matter we inhabit. But, through his collaboration with the physician Thomas Sydenham, he came to argue against the possibility of knowing anything concerning the true causes of disease, eventually denying that the study of anatomy could tell us anything about illnesses and how to cure them.[18]

[16] The *realist* does not think that the difference is between good and bad predicates, but between real (natural-kind) predicates and artificial ones. Normativity comes in, as it did in Plato and Aristotle, with the intensional structure that such a distinction engenders.

[17] See Specht 2009.

[18] See Walmsley 2008 for an excellent treatment of the relation between the two thinkers.

This radical rejection of causal explanations and scientific theorizing was an outlying view; indeed, it seemed to deny the tremendous success that Harvey's work had enjoyed all over Europe. Thus Locke was obliged to state clearly his reasons for holding it, and it is in these writings that the interconnection of pragmatism and nominalism that interests us took its characteristic modern form. Since it is impossible for us to know the inner causes of nature, it is impossible to establish a correspondence theory of truth, and the role of truth is taken over by something else, as Locke explains in his "De Arte Medica":

> …all speculations in this subject [the knowledg of natural bodys] however curious or refined or seeming profound and solid, if they teach not their followers to doe something either better or in a shorter and easier way than otherwise they could, or else lead them to the discovery of some new and usefull invention, deserve not the name of knowledg, or soe much as the wast time of our idle howers to be throwne away upon such empty idle phylosophy.[19]

That is to say, knowledge is to be called "true" not because it corresponds to the inner workings of things, whether these be souls or mechanisms, but only to the degree it allows human beings to flourish. As Locke's own career as an amateur physician flourished, so did his commitment to Sydenham's pious anti-realism, and they doubled down on this view in their collaboration on Locke's "Anatomia":

> Others of them have more pompously and speciously prosecuted the promoting of this art by searching into the bowels of dead and living creatures…to find out the seeds of discharging them, but with how little success such endeavours have been and are like to be attended, I shall here in some measure make appear.[20]

> All that anatomy can do is only to show us the gross and sensible parts of the body, or the vapid and dead juices, all which, after the most diligent search, will be no more able to direct a physician how to cure a disease than how to make a man; for, to remedy the defects of a part whose organical constitution, and that texture whereby it operates, he cannot possibly know, is alike hard as to make a part he knows not how is made.[21]

Locke was in fact what Kant later called a "transcendent realist" – his empiricism did not compel him to deny all structure to the mind-independent world, and he accepted much of Descartes's mechanical physics. But, like many

[19] In Dewhurst 1966: 83.

[20] Sydenham, handwritten remark on Locke's "Anatomia", in Dewhurst 1966: 85.

[21] Locke, introduction to "Anatomia", in Dewhurst 1966: 85.

Protestant fundamentalists up to the present day, he did not assign much importance to such a theory of stuff. Mechanical physics is useful to human beings who wish to manipulate lifeless Cartesian matter. But this is a pale shadow of the Aristotelian universe, with its natures and causes. Locke and Sydenham warn us above not to confuse such useful investigations of the material universe with an explanation of the true natures of things. On the contrary, all such explanations "do work" and "do not idle" only under the direction of our pragmatic goals.

The result is a peculiar inverted ontology, but one which now makes perfect sense: the *further* an object is from the network of human needs, the *less* we can know about its real properties. Spirits, including the Cartesian souls of other people, are out there, but are so infinitesimally *small* as to escape observation:

> …it is certainly some thing more subtile & fine then what our senses can take cognisance of that is the cause of the disease, they are the invisible & insensible spirits that governe preserve & disorder the aeconomie of the body.[22]

But this aperture on the mind-independent world shrinks to a vanishing point in later authors such as Hume. The latter does occasionally qualify his skepticism by observing, for instance, that necessary causal relations in things might well exist, but this does not change the fact that, even if they are out there, we still have no means of discovering them. At the limit, the theory of matter deserves the name of knowledge *only* insofar as it helps us achieve our practical goals.

By contrast, it is much easier to identify real kinds within humanity than it is to identify the real kinds of matter. Hume and his contemporaries remain realists concerning *human kinds*, a topic actively discussed throughout 17th and 18th c. Europe, which conviction becomes explicit when Hume discusses slavery:

> I am apt to suspect the negroes and in general all other species of men…to be naturally inferior to the whites. There never was a civilized nation of any other complexion than white, nor even any individual eminent either in action or speculation. …. Such a uniform and constant difference could not happen, in so many countries and ages, if nature had not made an original distinction betwixt these breeds of men.[23]

Like Locke, Hume also believes that regularities in comportment and ability point to real differences between kinds of human beings, that these differences are intrinsic and were implanted by God. But in his works on the foundations

[22] Locke, "Anatomia", in Dewhurst 1966: 91.

[23] Hume's "Of National Characters" as quoted in Immerwahr 1992: 481-482.

of natural science this realism is entirely offset by Locke's and Sydenham's arguments concerning the unknowability of the actual causal structure of these natures, meaning that, on those rare occasions where Hume does oppose slavery, he does so on purely *economic* grounds. These philosophies contain, and not by accident, both the materials for denying human rights to subjugated peoples, and those required to shed doubt any scientific claim that does not lead to greater riches, including the results of intelligence tests. The suspicion of essential inferiority can never be removed, since it is based on metaphysical convictions which are, on methodological grounds, beyond the reach of any experiment.

So in these authors, nominalist inductivism in its epistemological form makes all universal knowledge, and all universal laws, provisional and subject to doubt; however, in contrast to Berkeley, Kant and 20th c. empiricists, this does not mean that there are no real kinds in nature. Furthermore, neither Locke nor Hume denied that there were certain sciences, for instance arithmetic and algebra, whose clarity and distinctness made them impervious to such doubt. It was only in the case of geometry that the same worries occur as in the case of natural kinds, for geometry also required abstract ideas, or "transcendental forms" to mediate its inferences. Nineteenth-century nominalism, such as that found in the work of F.A. Lange,[24] and those influenced by him, such as Quine's teacher Peirce, results once the transcendental forms are again stripped out. This became the position of "naturalized Kantianism" that feeds into many of the authors discussed in the first sections of this paper.

10. *20th century pragmatism and HPS*

Quite obviously, one could object to this entire project on political grounds, arguing that this is simply the ontology you would expect from fundamentalist slave-traders, and that this explains its ongoing popularity in those part of the English-speaking world that were actively involved in that trade. For, realism about human kinds justifies our *not* extending full political rights to all human beings – merely to those capable of full rationality – and this barrier is absolute, grounded in essence; while realism concerning fundamental physics is either rejected or subject to endless skeptical doubt, because it threatens to produce a theory of the world that undermines the revealed truths of religion. But this is not a paper in practical philosophy, nor is it helpful to attack the doctrine on that field. These political consequences are seamlessly derivable

[24] I am deeply indebted here to Samuel Descarreaux, whose work on Lange has made clear the enormous impact he had on authors as diverse as Cohen, Peirce and Nietzsche.

from what appears to be "epistemology", even if, as we have just seen, that epistemology was always dependent on a radical nominalist metaphysics. Philosophers of biology working in this tradition have recently been forcibly reacquainted with that history. Since empiricism, including the logical empiricism of the Vienna Circle, endorsed Humean causality in this tradition, and defined the latter in terms of universalized material conditionals, it has found itself increasingly unable to find any explanatory value in Darwin's theory. To do so requires translating the extensional definition of "more fit" into some kind of causal explanation. Jerry Fodor's explicitly Goodmanian arguments lead to the usual result: causal explanations either fall prey to their disjunctivized twins, or they are tautological. The conclusion – that Darwin's theory is empty – was programmed already in the 18th c., by the ancestors of those who attack biological science today.[25]

11. *Goodmanian entrenchment*

Let us now turn to Goodman's own, nominalist solution to his paradox. Few pragmatists or Quinean naturalists have followed in his footsteps, because the task is enormous: a pragmatic reconstruction of a foundational science doubles the work. For, if the pragmatists are to provide a genuine alternative, they must embed the science in question in some larger utilitarian frame, in order to define concepts such as "truth", "consequence", "real predicate" in terms of utility. But then they must return to the original system to show the genuine theoretical consequences of the reconstruction. Goodman, to his credit, tried to do this, and the way his project ran aground is instructive precisely because it highlights the enormous task every pragmatist faces.

Goodman concludes his discussion of his New Riddle by *naming* what can count as an acceptable solution. The two coextensive predicates are distinguished by the fact that one is – in fact – *projected* by us, whereas the other is not. Thus what we are looking for is a theory of *projectability*. Since *projected* is a predicate that itself applies to *words*, statements about past projections are lists of observations, and *are* empirically verifiable. By contrast, a *projectable* predicate is one that *could*, and perhaps *should*, be used to make predictions concerning events which have not yet been observed, but which *will be* empirically verifiable. Thus, the task confronting us is to give criteria, referring only to actual past usage, which select "good" predicates for future use. This is the

[25] By contrast, parrying this attack inevitably involves invoking counterfactuals (Sober 2010: 606; Rosenberg 2013; Dubé 2019) and thus, if these are to be interpreted *objectively*, some form of causal realism.

problem of *projecting "projected"*, that is to say of defining a criterion, in terms of past events, which selects the "right" predicate to apply to future events.

Goodman's solution is to use induction on the past use of names: what distinguishes "grue" from "green" is their *history.* A term that has been successfully used in many inductions in the past is *entrenched*, such that, for any two predicates, the one with the higher degree of entrenchment is more projectable. If we endorse this definition, we make an empirical claim: the predicates on which we should induce are the ones that have been found, in the past, to have been repeatedly involved in successful, non-trivial inductions – these *will work better* in the future. It is in this sense very close to earlier definitions of "natural kind" in the work of philosophers such as Whewell and Mill. The classical metaphysician would have argued that, by means of experiments under varying conditions, we determine causal properties, thereby identifying species of objects with essential natures – horses, hydrogen atoms, top quarks, etc. But this relation can be inverted. We argue instead that we *call* those classes of things "natural kinds" whose predicates have often been used in describing successful experiments that we took to have confirmed universally quantified propositions. The proposal is that, faced with a deadlock, we select the one that worked well in the past.

Goodman should be commended for attempting what most others in this tradition only gesture towards. He truly does reconstruct the notion of a natural kind within a larger formal system that meets the scruples of the radical inductivist, quantifying only over past events that actually occurred. But he nevertheless failed in his attempt, and these failures are instructive because they point at the core weakness of the entire nominalist-inductivist tradition. So I will conclude this discussion by briefly summarizing the key difficulty, after which we will return to the constitutive role of this work in late-20th c. HPS, in the work of historical epistemologists such as Kuhn and Hacking.

12. *Entrenchment of kinds and laws*

On Goodman's approach, higher "entrenchment" accrues to predicates that have been used a lot – *green* has history, *grue* does not. But "use" is *too weak* for this definition to do any work. For instance, "delivered by Santa Claus" is used a lot, but is not predictively successful. So "use" must be restricted, and the definition strengthened: "used" shall mean "mentioned in many universally quantified propositions that were successfully used to make predictions in the past".

But now the definition is *too strong.* Consider, for instance, the periodic table of the elements. The projectability of the names that appear in it is not

just a function of the fact that they're used a lot. Some, like "Seaborgium", refer to substances which have rarely existed in the history of the universe, and have equally rarely, if ever, been used to make predictions. Goodman tries to solve this problem by bootstrapping. Entrenchment of low-level predicates, e.g. *Carbon*, confers projectability on some higher-level ones, e.g. the periodic table itself, and this then "trickles down" to *Seaborgium*. Put otherwise, Goodman must construct nominalised higher-level genera (replacing *natural genera)*, in order to handle species-concepts with short entrenchment-histories. And this recurs across all concept-types, since new scientific concepts are typically introduced in the vocabulary of some prior science, which in the mature physical sciences will generally involve some "laws" that must be interpreted nominalistically as well.

At this point, Goodman's theory collapses under its own weight, as he tries to introduce principles to disseminate entrenchment within the conceptual network. Roughly, *Carbon* is a scientific concept that inherited the deep entrenchment of words such as *coal*, and, perhaps still earlier, *wood* and *burn*. This entrenchment should contribute to that of higher-level species-concepts such as "element of the periodic table", which will then percolate to Seaborgium and other such problematic concepts. While *FFF* does not treat of mathematics, the nominalist-inductivist position here is of long standing: mathematical propositions are highly purified empirical propositions, whose pragmatic value derives from their use in the empirical sciences. Considered on their own, they have the status of dictionary definitions (conventionalism) or are simply imprecise, because oversimplified, physical sciences.

We can diagnose the problem with reference to its historical origin, and the entrenchment of that view. When Sydenham and Locke denied explanatory value to anatomy, they committed themselves to diagnoses based on correlations between observables on the *surface* of the organism, while denying any importance to its internal mechanisms. That same conviction, applied in physics, leads to the doctrine that one should not speculate concerning the nature of *forces*, for these are fully describable with reference to their kinematic effects. But each time causes, microstructure and mathematics are called into doubt, something must be introduced into the nominalist theory to replace them, in the same way that utility replaced truth. At the limit, one attempts to explain the economy and utility of laws of nature by claiming that *systematization* allows us "to doe something either better or in a shorter and easier way".

But no one actually *believes* that this is why we have a periodic table, or that atoms are fictions, except of course the pragmatists themselves. Since the pragmatist holds that "real" means at best "entrenched", he must respond to this fact: Doesn't the entrenchment of this realist conviction, or for that matter the

Platonist conviction of mathematicians that their abstract objects exist, have equal claim to validity as does the proposal of the pragmatist himself? That is to say, once it is conceded, on pragmatic grounds, that realism and Platonism are heuristically useful fictions, and that utility is the only criterion of reality that matters, why not *prefer* them to the nominalist's story, which makes no stronger claim for *itself* than to be a useful fiction?

Goodman's failure is, from this point of view, the formal correlate of a systematic problem. Utility *can't* provide the criteria we need to understand theoretical science, because the problem with Seaborgium is general. Our economic interests are much less finely-grained than our theoretical knowledge. It is therefore not possible to explain the surplus detail of mathematical and natural scientific reasoning without either trivializing the latter, or suggesting that mathematicians and scientists are gravely deluded concerning their own ontological and methodological commitments. We see this failure in much recent pragmatically driven HPS, when it is suggested that the emergence of Einstein's theories can and should be explained with reference to railroad time in the 19th c. No one seriously attempts this with mathematics any longer, which is one reason that Anglo-American history of mathematics is dying out.

The worry is that, in the case of mathematics, pragmatism has been *empirically* falsified – it is simply not the case that the grounds of mathematical justification lie in empirical science, not even in the case of applied mathematics. When we visit their department, we find mathematicians justifying by means of apodictic proof. It does no good to respond that (some of) this mathematics may eventually be applied by empirical scientists, nor even that mathematicians now make use of computer assistance in generating some of their proofs. For in neither case is the standard of validity the practitioners observe empirical and inductive. Should the pragmatist concede this, all while arguing that the mathematicians will enjoy more "success" if they adopt such standards instead of the ones they use in fact, then that is a *normative proposal*, and not an induction on past observations,[26] for the history of the field suggests the exact opposite.[27]

[26] One should carefully distinguish between the claim that *mathematicians would enjoy greater success if they collaborated with empirical scientists* and the far stronger claim being rejected: that *they would enjoy greater success if they followed empirical and inductive methods within mathematics itself*. It is characteristic of the tradition being discussed to fall back on the first sort of claim, which most people would accept, when challenged on the second, radical claim that is the real content of their theory.

[27] To take just two examples among many: the long and intricate development of Riemannian geometry took place in the absence of any pragmatic need from the physical side; mathematicians such as Hilbert, who laid enormous emphasis on connecting work in mathematics to the needs and aims of physical science, always viewed their contribution as *foundational*. In axiomatizing scientific

In the case of the philosophy of the empirical sciences, the original connection between pragmatism and religious faith has reasserted itself: nominalist arguments such as those of Fodor are now being used to undermine the foundations of evolutionary biology. Our short historical epistemology of "historical epistemology" show us why and how this is possible. The first use to which Locke and Sydenham put their empiricist philosophy – the rejection of physiological and anatomical *explanations* of disease – was, from the beginning, a rejection of causal explanations in biology. It should not come as a surprise that it has exactly the same consequences today.

13. *Projecting "historical epistemology"*

The successful, because formally irrefutable, part of *FFF* is a demonstration that unconstrained nominalism leads to the result that all (non-falsified) hypotheses are equally credible, in other words that an extensionalist theory of induction is impossible. Since induction was to be the *only* means of generating general knowledge for Quine, including mathematical knowledge, this result is terminal. But Goodman also does not succeed in answering his New Riddle, at least not on terms that he can accept. Since he has strong reasons for doubting that causal relations and natural kinds, should they in fact exist, could be *known* otherwise than by induction on actual past observations, he agrees with Quine in rejecting all realist solutions as non-responsive, and extends this ban to *possibilia*. In consequence, the indisputable scientific role of such concepts and forms of reasoning requires a substitute, and this must be constructed within linguistics. If we set aside the significant obstacles to carrying out this programme positively, we get an influential *negative* doctrine. There is no such thing as *neutral* inductive verification: in every induction, thus in all empirical science, historical factors intervene in the form of our inherited conceptual scheme. In studying the history of these schemes, we do "historical" epistemology.

Goodman himself acknowledges that this result is in part a Kantian one, and we can now see why. Like others in this tradition, such as Husserl, he ascribes an essential cognitive role to a *prior* scheme, and since he is an empiricist epistemologist, this priority is temporal *only*. The epistemologists who follow in this tradition – above all Kuhn and Hacking – are therefore *historical* epistemologists. By identifying the paradigms, disciplinary matrices and

theories, we aim to isolate and clearly distinguish the empirical and "synthetic" parts of a theory from its logico-mathematical presuppositions. Maybe Hilbert was wrong to think that way, but it is a matter of fact that he did, and it can scarcely be denied that his "programme" was one of the most fruitful projects of our recent mathematical past.

styles of past scientists, we are not *merely* telling factual history; rather, we are inquiring into the conceptual presuppositions of our present, and therefore also sketching in our future. In opposition to thinkers in the French and German traditions, the "conceptual schemes" of pragmatists and naturalists are historical facts, but this difference is less than it might appear. For while thinkers such as Canguilhem roundly criticized Kuhn on just this point, he could not articulate his objection in terms that anyone working in that tradition would accept or understand. While Goodman had himself already put his finger on the key problem: counterfactual conditionals and *possibilia*, which, like a bump in the rug, stubbornly resist being smoothed away. This is why, slowly but surely, younger philosophers have moved towards variants of metaphysical realism.

For, as should now be clear, the *function* of conceptual schemes within the nominalist-empiricist tradition is much the same as it is in transcendental versions, and indeed within metaphysical realism. As the nominalist reconstructs the concepts of the sciences, he embeds them within a socio-linguistic structure – we saw how Goodman replaced the concept of a natural kind with that of linguistic entrenchment, which can now function as the *explanans* of statements such as "this stone would have been green even if I hadn't examined it." Given that entrenched predicates have a good – which simply means "long and successful" – epistemological history of projection, it is natural that speakers would say things like this, for a *habit* has formed in them, or in their ancestors. Nor should we be worried if they express this fact poetically, in terms of realism, claiming that "green is a property, and no mere predicate, and things have their properties whether or not we examine them." For this is just a complex way of expressing the deep entrenchment of the term. If one objects that this entrenchment itself is better explained with reference to the causal powers of natural kinds, the nominalist will respond with a *fact*: those natural kind concepts have worked well in the past, and this history has left a trace in us, or in our ancestors. This habit, which is perhaps by now entrenched in our neurology, is all we need to explain our current preference for them. Such an explanation is supposed to be *natural-scientific*, uncontaminated by metaphysical, or foundationalist superstitions. But is it be "epistemology"?

I think that we must answer in the negative. For, little has changed since Chomsky's devastating critique of linguistic behaviourism. Concepts such as "logical inference", "true", "justified" and "belief" are unquestionably part of the justificatory practice of empirical scientists, and the disciplines that study them have long been called "epistemology". It is entirely possible to study the users of that framework empirically, just as it is possible to observe

the brain of a logician who is executing a proof. But there are bad scientists, and poor logicians – choices will have to be made by the researchers regarding whose brains, and which neural processes are to count as good ones.[28] The psychologist does not have to make that distinction, until he claims that he is doing epistemology – then he owes us such a definition. And here it does no good to be told that the good ones are the ones that lead to more utility. For this is, when all is said and done, exactly as helpful as being told that the best investment strategies are the ones that lead to the highest returns, or that the best shoes are the ones that are most durable and comfortable. The craftsman knows this already, and still it tells him next to nothing about *how to make good shoes*.

When it comes to *historical* epistemology, the situation is fundamentally no different. The normative versions all involve a transcendental residue: *epistémès*, historical a prioris. Their naturalized cousins contain naturalized copies of these: disciplinary matrices, paradigms, language-games, styles. The first type of theory is vulnerable because it cannot explain the ontological status of these transcendental residues. The second avoids that problem precisely because the structures in question are naturalized; however, for that very reason, they have no normative force. The nominalist-inductivists themselves have already clearly shown us why. On their own account, such investigations say nothing definite about what *should* happen after time *t*. And while it remains true that now, at *t*, we do have certain *inclinations* about how to proceed, that was also never in doubt. The problem is that different people have different inclinations, just as methods that have worked well in the past often do not do so going forward. In consequence, while our history has much to teach us here, we can project its lessons on the future only to the extent that we have analyzed it normatively to begin with.

David Hyder
dhyder@uottawa.ca
University of Ottawa

[28] Put simply, the question whether the logician's proof is *valid* cannot be answered by studying her brain, nor by examining the history of logic. This does not mean that such examinations are without interest or utility, merely that they do not and cannot on their own provide an answer to the *logician's* question. See Kim 1988: 391ff., as well as footnote 26 of this paper.

References

Cohnitz, Daniel and M. Rossberg, 2006, *Nelson Goodman*, Acumen, London.

Declos, Alexandre, 2019, "L'énigme du 'vleu' et l'hyper-nominalisme de Goodman", in *Igitur*, 10, 1: 1-27.

Dewhurst, Kenneth, 1966, *Dr. Thomas Sydenham (1624–89): His Life and Original Writings*, Wellcome Trust, London.

Dickson, Michael, 2004, "Quantum Reference Frames in the Context of EPR", in *Philosophy of Science*, *(Proceedings)*, 71, 5: 655-668.

Dubé, Jean-Philippe, 2019, *Causalité dans la théorie darwinienne*, University of Ottawa, MA Thesis, DOI: dx.doi.org/10.20381/ruor-22956.

Einstein, A., B. Podolsky, and N. Rosen, 1935, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", in *Physical Review*, 4, 1: 777-780.

Goodman, Nelson, W.V. Quine, 1947, "Steps Toward a Constructive Nominalism", in *Journal of Symbolic Logic*, 12, 4: 105-122.

Goodman, Nelson, *Fact, Fiction, and Forecast*, 1979, 4th ed., Harvard University Press, Cambridge MA.

Hacking, Ian, 1993, *Le plus pur nominalisme: l'énigme de Goodman: «vleu» et usages de «vleu»*; Fr. tr. By R. Pouivet, Éditions de l'éclat, Paris.

Hyder, David, 2003, "Foucault, Cavaillès, and Husserl on the Historical Epistemology of the Sciences", in *Perspectives on Science* 11, 1: 107-129.

Immerwahr, John, 1992, "Hume's Revised Racism", in *Journal of the History of Ideas*, 53, 3: 481-486.

Kim, Jaegwon, 1988, "What Is 'Naturalized Epistemology?'", in *Philosophical Perspectives. Vol. 2, Epistemology*: 381-405.

Rosenberg, Alexander, 2013, "How Jerry Fodor slid down the slippery slope to Anti-Darwinism, and how we can avoid the same fate", in *European Journal for Philosophy of Science*, 1, 3: 1-17.

Sober, Elliott, 2010, "Natural Selection, Causality, and Laws: What Fodor and Piatelli-Palmarini Got Wrong", in *Philosophy of Science*, 77, 4: 564-607.

Specht, Rainer, 2009, "Experience and Hypotheses: Opinions within Locke's Realm", in M. Heidelberger, and G. Schiemann, eds., *The Significance of the Hypothetical in the Natural Sciences*, De Gruyter, Berlin and New York: 39-58.

Walmsley, Jonathan, 2008, "Sydenham and the development of Locke's Natural Philosophy", in *British Journal for the History of Philosophy*, 16, 1: 65-83.

Ian Hacking's metahistory of science

Manolis Simos and Theodore Arabatzis

Abstract: In this paper we attempt a critical appraisal of the relation between history of science and philosophy of science in Ian Hacking's styles of scientific reasoning project. In our analysis, we employ a distinction between "historical *philosophy* of science" and "philosophical *history* of science": the former aims at addressing philosophical issues, while the latter aims at telling stories about the scientific past that are informed by philosophical considerations. We argue that Hacking practices historical *philosophy* of science; discuss how his approach is differentiated from the so-called confrontation model; and show that he opts for a strong integration between history and philosophy of science. Finally, we discuss the historiographical implications of his approach and suggest that his aim to maintain a middle position, on the one hand, between contingency and inevitabilism, and, on the other, between internalism and externalism in the explanation of the stability of scientific knowledge, is compromised by his philosophical commitments.

Keywords: Ian Hacking; styles of scientific reasoning; history and philosophy of science; contingency/inevitability; internalism/externalism

1. *Introduction*

This paper aims at a critical appraisal of the relation between history of science and philosophy of science in Ian Hacking's work. Specifically, we focus on the series of essays that comprise his *styles of scientific reasoning project* and examine his theoretical reflections on that relation. In these essays, Hacking provides us with elaborate metahistorical reflections: theoretical reflections on the relation between philosophy and history, and on his own philosophical and historical practice, that is, on his own way of practicing philosophy and history, or, better, philosophy *cum* history.

Our approach to these reflections is guided by three closely interrelated issues. First, we focus on Hacking's theoretical stance on the relation between history of science and philosophy of science. This question takes the form of an inquiry concerning whether Hacking's approach constitutes a 'historical *philosophy* of science' or 'a philosophical *history* of science'. The former aims at

ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.314

elucidating or taking a stance on philosophical issues by drawing upon historical material, while the latter aims at telling stories about the scientific past that are informed by conceptual and philosophical considerations.[1] In discussing this question, providing textual evidence and interpretative justification, we argue that Hacking self-reflectively practices 'historical *philosophy* of science'; that is, he articulates a philosophical stance in response to philosophical issues and he argues for it historically.[2]

Second, we argue that Hacking's approach is differentiated from the so-called *confrontation* model in history and philosophy of science (HPS), and we sketch the alternative integrated approach he suggests. As we discuss in the next section, after the critique of logical positivism in the 1960s and 1970s, the resulting turn to a naturalized philosophy of science involved a specific idea of combining history and philosophy of science. According to that idea, history becomes an *ancilla philosophiae*, providing empirical data for the (dis)confirmation of philosophical theories about the nature of science.[3]

Third, given that Hacking's approach is more the one of a philosopher employing a historical method, than that of a philosophically sensitive historian, we discuss the historiographical implications of Hacking's approach, which aims at maintaining a middle position concerning fundamental historiographical issues, such as contingency and inevitabilism, and internalist and externalist explanations of the stability of scientific theories. We show that his approach compromises the middle position he intends to adopt.

To that effect, our argument is developed in three sections. In the next, second section, we present the framework within which our three issues are approached. In the first part of the third section (3.1), we present the historiographical structure of Hacking's styles approach, that is, we present a schema that illustrates the way he uses history. This schema shows that Hacking's styles project constitutes a historical *philosophy* of science, and that he adopts a strong integrated approach concerning the relation between history and philosophy of science. In the second part of the third section (3.2), we focus on the question of stabilization of styles. Here, we substantiate in a concrete manner the way Hacking uses history in order to argue for his philosophical claims, and we critically discuss the historiographical implications of his approach.

[1] See Arabatzis 2017.

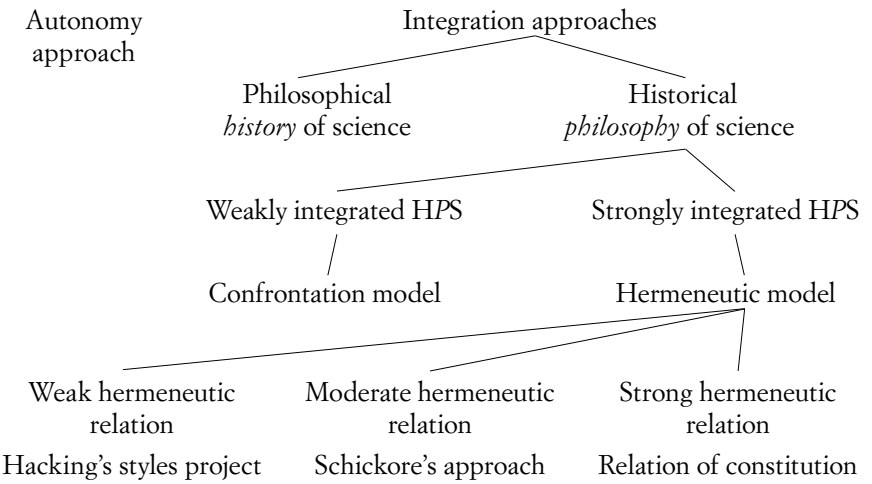
[2] For a different perspective on how Hacking understands the relation between history and philosophy, see Kindi 2014.

[3] See Schickore 2011: 466, 477, and *passim*.

2. *A metahistorical framework: Mapping the relations between history and philosophy of science*

In this section, we will present a framework within which the three above issues will be discussed. This framework can be represented by the following schema that maps some possible relations between history and philosophy of science, registering a provisional placement of Hacking's approach in it.

Fig. 1 – Some possible relations between history and philosophy of science.



According to the autonomy approach, philosophy and history of science are ontologically and epistemologically distinct and autonomous. They have different subject matters and adopt different methodological approaches, while, more importantly, neither their subject matter nor their methodological approach necessitate an interaction. The underpinning analogy here is between scientific inquiry, say chemistry, and the history of scientific inquiry, that is, the history of chemistry. Thus, philosophy of science is modelled upon science itself. It constitutes a metascientific inquiry, whose subject matter is science itself. In contrast, history of science falls under the *Geisteswissenschaften*. It constitutes an interpretative enterprise of matters past, pertinent to scientific inquiry and practice simpliciter. Philosophy of science is ahistorical; history of science is unphilosophical.[4]

We agree with Schickore that the autonomy approach is conceptually

[4] See Schickore 2011: 461.

flawed. On the one hand, philosophy of science has an inbuilt historical component. The construction of a philosophical theory about science, including its development, requires interpretative work. It involves a reconstruction of at least the recent past of specific scientific fields. In that sense, it cannot be completely ahistorical. On the other hand, the converse holds for history of science. The construction of a narrative account of, say, a specific scientific episode – to take a near uncontroversial, if not always accurate, description of what historians do – involves philosophical presuppositions and philosophically laden categories. The individuation of a historical case as a "scientific episode", its relation with other events within and beyond "science", and its falling under a particular meta-scientific category (e.g., "discovery" or "experiment") are only a few representative ones.[5]

Integration approaches to the relation between history and philosophy of science take into account the preceding critical points. The confrontation model can be considered the most representative and dominant version of what we can call 'weak integration approaches'. It acknowledges the relative autonomy of both fields, yet allows a weak interaction between them. According to the distinction we introduced above, this model constitutes a 'historical *philosophy* of science'. Philosophy of science, according to this model, has an ontological and epistemological primacy over the history of science. Philosophy of science formulates theories or theses on, mainly, scientific change, progress, and rationality, which "had to be *confronted* with historical episodes. The role of history was to provide the data for the evaluation of philosophical theories about science" (Schickore 2011: 464).[6] Thus, history of science assumes a merely ancillary role in philosophical reflection about science, without its being constitutive of the philosophical issues under investigation.

This conception of the relation between historical data and philosophical hypotheses encounters several problems: as philosophical theories that seek to be (dis)confirmed by the available (historical) data seem to be modelled upon scientific theories, they face all the well-known problems associated with the testing of scientific theories. The first problem is the theory-ladenness of the historical (*qua* empirical) data. In the same way that there can be no neutral empirical data for the appraisal of scientific theories, there can be no neutral historical data for the evaluation of philosophical theories. Second, the very method of theory appraisal to be followed has to be specified and justified. For instance, as Lakatos pointed out, in evaluating a philosophical theory of scientific change one can adopt an inductivist approach, or a Popperian falsifi-

[5] Cf. Arabatzis 2006; 2012; 2017.

[6] Emphasis added. Cf. Vagelli 2019.

cationist approach, or finally a Lakatosian metamethodology.[7] Third, the privileged status of *historical* material over other kinds of *empirical* evidence has to be equally justified. According to Schickore, "historical study becomes just one option among many other empirical approaches, such as sociology, cultural studies, cognitive science, ethnography, anthropology, media studies, and so on" (Schickore 2011: 470).[8] Fourth, a final problem is the underdetermination of philosophical (*qua* scientific) theories by historical (*qua* empirical) evidence. Namely, it is possible that the same historical evidence can justify different – or even opposing – theories about science.[9]

The critique of the weakly integrated historical *philosophy* of science and its most representative version, the confrontation model, may lead to a stronger integration between history and philosophy of science. Schickore's account, succinctly summarised in her diagnosis that "[*u*]*nderstanding science results from a hermeneutic procedure*, in which preliminary concepts and frameworks and initial case judgments are *modified* and *adjusted* until a cogent account is obtained," is differentiated from the weakly integrated historical *philosophy* of science in two ways (*ibid*.: 478, emphasis added).

First, as we said earlier, her approach suggests an inextricable relationship between history and philosophy of science. A philosophical apparatus or toolbox of more or less specified "concepts", "frameworks", and "judgments" is suggested, in light of which the historical record is approached. However, this provisional apparatus is open to revision in light of the historical material encountered. This iterative process encapsulates the hermeneutic quality of the approach.

Second, following Hacking following Foucault, Schickore's invoking of the notion of the 'history of the present' as something that "should remain part and parcel of our present efforts to understand the sciences" can be understood in a double way (*ibid*.). First, her approach is differentiated from weak integration approaches in that the historical aspect of the enterprise is not considered a necessary addendum – something to which the advocates of the autonomy approach would eventually concede, realising that their inquiries inevitably involve the reconstruction of, at least, the very recent past instances of their subject matter. Rather, history is seen as an integral part of philosophical understanding. Second, and most important, as we will see in the next sections

[7] See Lakatos 1980.

[8] It could be argued that Hacking's invoking of other disciplines such as anthropology, sociology, and cognitive science reinforces Schickore's point. As it will be shown below, though, Hacking does not consider history methodologically replaceable by those disciplines.

[9] Schickore (2011) does not refer to this problem. It derives though from the analogy between philosophical theories and scientific theories in the way described.

(where we discuss Hacking's work), the notion of the history of the present amounts to a philosophical investigation of a contemporary (meta-)scientific entity, such as objectivity. The study of the historical emergence and development of such an entity is constitutive of its philosophical understanding.

Building on Schickore's distinction, we can draw an additional one, elaborating on her schema. The hermeneutic model can appear in three versions: weak, moderate, and strong (see Fig. 1). In the moderate version, the historical material retains an ontological autonomy. The philosophical apparatus orders the historical record and singles out the facts, but it does not constitute them. Moreover, philosophical concepts get revised in light of the historical record. Thus, in this version – which seems to be Schickore's own – the distinction between a 'historical *philosophy* of science' and a 'philosophical *history* of science' seems to break down in favour of a seamless HPS approach.

The strong version is more radical. The relation between the philosophical claims and the historical material is such that historical facts can be identified only within and in virtue of a philosophical perspective.[10] This version brings to the fore the issue of the justification and validity of the hermeneutic approach. The thinkers who adopt it are either indifferent to this issue – embracing thus a strong relativism – , or reduce the justification of their approach to its contingent reception.[11]

Finally, Hacking seems to follow a weak version of the hermeneutic model, by retaining the element of bilateral revisability. Yet, as we will see, he prioritizes an overarching philosophical perspective. This provides the framework within which the hermeneutic interplay between the historical record, on the one hand, and the philosophical ideas, on the other, takes place.

3. *Hacking's styles project*

3.1. Styles as historical philosophy of science

3.1.1. Styles as historical philosophy of science: the structure of the relation

Relations between the history and the philosophy of the sciences are often debated and sometimes contested. My interest here is collaboration. I shall describe a new analytical tool that can be used by historians and by philosophers for different purposes.

[10] This idea is endorsed by two thinkers that Hacking explicitly draws upon, namely, Foucault and Latour.

[11] See, for example, Latour and Woolgar 1986: 257.

It is a specialized, indeed technical, version of an idea often used or abused elsewhere: "style." […] The two uses, by historians and philosophers, are complementary but to some extent asymmetric. The historian may conclude that the philosopher's use of the tool is bunk, irrelevant to understanding the past. But the philosopher needs the history, for if the tool does not provide a coherent and enlightening ordering of the record, then it has no more place in sound philosophy than would any other fantasy (Hacking 2002c: 178).

Hacking's pithy description can be unpacked in a series of points. First, its importance is underscored by the produced rhetorical effect: the above lines introduce his paper. Second, although Hacking assumes the standard distinction between history and philosophy of science, qualifying their relations as 'asymmetric', he explicitly suggests an integrated HPS approach. Third, he refers to the historian's task in terms of philosophical history of science, while, when he turns to the philosopher, he switches to historical philosophy of science. Fourth, and most important, in the latter case, the use of history is essential for the philosophical enterprise. According to Hacking's own telling formulation, without historical input philosophy cannot be sound, where soundness is related to some sort of empirical anchoring. A philosophical claim – let alone theory – that cannot be related to the historical record is excluded: Hacking does not talk about a philosophical error, or about the possibility of corroboration of a philosophical theory at a further point in time. Rather, the very *aim* of philosophy of science is to "provide a *coherent* and *enlightening ordering* of the record" (emphasis added). As such, the difference from the confrontation model is evident. In light of our comments in the previous section, an "ordering of the record" would be at odds with the confrontation model, while Hacking's formulation explicitly envisages the historical record as not coming in the form of a pre-packaged set of claims ready to be compared directly with a theoretical framework.

Hacking's latest remarks on the matter both corroborate and deepen the above interpretative points:

The styles project uses the past as a way to understand the present. Although it has suggested historical research to others, and draws on far more historical data than it cites, in itself it adds no new content to the history of science. The accounts of the past to which it refers are (disconcertingly for many readers) as often folklore as archive-based research. Anthropology, sociology, and cognitive science, especially of the more speculative sort, are also invoked. In short, the project is philosophy attentive to, but not awed by, many neighbouring bodies of knowledge and theorizing. (Hacking 2012: 600)

First, again, Hacking describes explicitly his project as a philosophical one. In fact, to that effect, he corrects his earlier view: "The 1992 title, '"Style"

for historians and philosophers', was a mistake, for the paper addressed philosophy, not history" (*ibid.*: 601). Moreover, the philosophical aspect is further underscored. Hacking is explicit that his philosophical approach – despite its opening a space for historical research – is not in itself a contribution to the history of science.

Second, as we will elaborate further, Hacking's styles are situated within his attempt to address rather traditional philosophical issues: scientific rationality, method and truth. His philosophical agenda is to maintain a critical distance between two poles: on the one hand, a metaphysical and epistemological realism, and, on the other, a relativism amounting to subjective idealism, in which scientific truth and method become a matter of *ad libitum* decisions (Hacking 2002c: 196). However, given his rejection of ahistorical philosophy of science, the use of history becomes necessary. As it will become more apparent below, history acquires a double role: it is needed to justify the critical distance maintained from both of these two poles.

Third, more specifically, this middle position can be further understood in terms of the notion of the history of the present: "The history that I want is the history of the present. That is Michel Foucault's phrase, implying that we recognize and distinguish historical objects in order to illumine our own predicaments" (*ibid.*: 182). Thus, Hacking aims at an understanding of the present condition. And given that this condition is the product of historical developments, its understanding cannot but be at the same time a historical one. Moreover, this understanding of the present is philosophical, since styles are the conditions of possibility of our contemporary condition. Styles provide "an account of how conceptions of objective knowledge have come into being", and the link between Kant and the history of the present becomes explicit (*ibid.*: 198).

The previous three points provide the structure of Hacking's general philosophical position. Furthermore, the summarizing extract above raises three new important points concerning the very use of history in the structure just outlined. First, Hacking acknowledges that his philosophical project derives from and is supported by historical data, even if these are not explicitly mentioned. Second, the use of the historical record is not exclusive. Rather, the use of other empirical fields is acknowledged. Moreover, he does not prescribe any kind of hierarchy, according to which a specific field of inquiry and body of data would acquire ontological, epistemic, or explanatory priority. Third, Hacking distinguishes between two types of historical accounts: "folklore" and "archive-based". This distinction does not primarily concern the evaluation, acceptance status, or degree of entrenchment that some historical interpretations have within the community of historians. Rather, we take it more to distinguish between historical data proper and historiographical ideas, which

frame the historical data. "Folklore accounts" refer to the latter. Hacking describes them as "familiar legend[s]" and "popular myths of origin", associated with Galileo, Boyle, and the air pump itself (*ibid.*: 185). They are used as metonymies for the crystalized state of a style (Hacking 2012: 607). As such, they can be understood as middle range historical abstractions around which historical evidence is mustered. They constitute, at the same time, both historical generalizations *qua* facts, and the perspectives from which historical material is approached.

These last three points document Hacking's differentiation from the confrontation model. His account differs from it in both 'horizontal' and 'vertical' terms, that is, concerning both the relation of history to other fields, and the relation between philosophical claims and historical evidence, respectively. As regards the first, in lieu of the prioritization of history and the uncritical bracketing of neighboring discourses that characterize the confrontation model, Hacking adopts a more inclusive approach. As regards the second, as we saw, historical evidence is not conceptualized as a relation of correspondence between factual statements and theoretical claims ready to be confirmed or disputed. Rather, for Hacking, historical evidence seems to come into bundles mined out from the archive – understood in a broad sense – by philosophical claims, historical abstractions, and historiographical considerations.

3.1.2. Styles as historical philosophy of science: a hermeneutic interpretation

This subsection discusses the second point that we would like to make, that is, the dynamic, coherentist, and hermeneutic aspect of Hacking's approach. This aspect can be expressed in three interrelated remarks. First, apart from the fact that Hacking describes his project in philosophical terms, he indicates that the categories he uses are open to modification, and lays out his project in terms of a process (Hacking 2002c: 182). Although the specific content of this process will be described in the next subsection (3.2), we can offer a schematic representation of its structure in terms of the following elements:

a) Hacking's adoption of a middle position, concerning realism and nominalism, inevitabilism and contingency, and internalism and externalism regarding the explanation of scientific stability;
b) his appropriation of Crombie's notion of styles (*ibid.*: 186);
c) his invoking of philosophical and historiographical tools, such as Foucault's notion of *epistemes*, Foucault's and Comte's notion of *positivity*, and conceptions of truth and meaning from analytic philosophy;
d) the application of this philosophically informed apparatus to Crombie's styles, resulting into Hacking's own account of styles (*ibid.*: 198).

Second, although Hacking uses the distinction between history and philosophy of science, and describes his project in philosophical terms and Crombie's in historical ones, in his conclusion the distinction and corresponding division of labor between the historian and the philosopher is almost suspended:

> For all the manifest differences of endeavor between the historian and the philosopher, they have this in common: we share a curiosity about our Western "scientific" vision of objectivity. […] Yet I would not push this division of labor too far. […] however much the historian may abjure philosophical issues, every sound history is imbued with philosophical concepts about human knowledge, nature, and our conception of it. And aside from central shared concerns, there is a more general predicament that the historian and the philosopher experience. Crombie was powerfully aware of the *reflexive* elements of his volumes. He knew that he who describes a certain vision of ourselves and our ecology has that vision himself. (Hacking 2002c: 199, emphasis added)

This is important for the following two reasons: first, a philosophical history of science (that is, Crombie's) suggests a strong interrelation between philosophical ideas and historical material. In other words, Crombie's account and Hacking's appropriation of it do not share the basic structure of the confrontation model. Second, and more important, the presentation of Crombie's styles in these terms indicates that Hacking does not just use a set of historical facts. Hacking's metaphors of 'legends', 'continents', 'waterfronts and piers' constitute the abstractions by means of which Crombie's account is reformulated. In that sense, we cannot talk about a philosophical use of historical data, but of philosophical elaborations of historiographical concepts, metahistorical ideas, narratives, and facts. Furthermore, we cannot talk about a relation between pure philosophical ideas (Hacking) and raw historical facts (Crombie), but rather between historically informed philosophical ideas, on the one hand, and philosophically laden facts, on the other.

Finally, third, Hacking observes that the philosophical relevance of Crombie's work, that is, of the efficacy of the tool Hacking selects to use, "is not a matter of principle", but is assessed by "the success of the resultant philosophical analysis" (*ibid*.: 186). In turn, this success is not assessed by an external set of criteria, but it depends largely on the success of the tool used in enabling a "coherent ordering and analysis of European scientific practice and vision" (*ibid*.: 198).

3.2. The stabilization of styles: historiographical remarks

We mentioned Hacking's attempt to maintain a critical distance from both poles of metaphysical *cum* epistemological realism and of radical relativism. In

light of what we have argued so far, this middle position can be unpacked as follows. We see deep historical changes in our most systematic, elaborate and sophisticated, collective interactions with the world. At the same time, these interactions are characterised by an irreducible stability. Crombie's notion of styles of scientific thinking, renamed by Hacking, initially, as styles of scientific reasoning, and, later, as styles of scientific thinking and doing, captures this characteristic stability of our systematic collective interactions. This idea is suggested explicitly by Hacking as an intersubjective, historicized, and pluralist version of Kant's conditions of possibility of truth and knowledge (*ibid.*: 181, 198). It is intersubjective *qua* collective, as these conditions of possibility do not refer to the structure of a transcendental subject; it is historicized, as styles are not atemporal but historically and locally specific; and it is pluralist, as styles amount to "disparate ways of thinking" (Hacking 2002b: 170). Thus, styles are conditions of possibility of objectivity, and these conditions of possibility are – roughly put, yet allowed by Hacking's own quips – historically determined nexuses of fundamental methodological practices.

As we saw, the notion of styles enables Hacking to adopt a position of mitigated relativism. Furthermore, it is this very notion that enables him to adopt a position of mitigated contingency, too:

> *There is no deep reason for, or cause of, the appearance* at different times of a few distinct genres of scientific inquiry, often detectable in ancient Greece, and still flourishing. […]. They began to stabilize but also continued to evolve in *an endless cycle of contingencies. This anarchic story is not quite that of a random walk, but there was no foreordained right route.* (Hacking 2012: 600, emphasis added)

Concerning the emergence, transformation, and demise of styles, Hacking holds explicitly a contingentist position (Hacking 2002c: 195). At the same time, he mitigates his position acknowledging constraints on contingency. First, the autonomy of a style of reasoning transcends the historical and social contingencies from which it emerged (*ibid.*: 196; 2012: 600). Second, Hacking's reference to Foucault is telling. Foucault's epistemes as conditions of possibility of the production of scientific statements have been criticized for their holistic, self-enclosed character:

> [R]ecall complaints addressed to Michel Foucault that he never explained why epistemes die out, in particular why his Renaissance episteme of resemblance expired. *I do not believe that one can give purely internal explanations of why we abandon certain practices, but have no confidence in external explanations either. It does not discredit the philosopher's use of styles of reasoning that it leads directly to such historical chestnuts; the contrary, I should imagine.* (Hacking 2002c: 195, emphasis added)

Here, the distinction between internal and external factors acquires a slightly different meaning compared to the standard one. The notion of external does not refer only to social factors – that is, external to scientific practices, traditionally considered –, but also to a kind of input from an extrastylistic reality.[12]

To discuss Hacking's use of history in his account of the techniques for stabilizing styles, we will employ a distinction that we find illuminating, namely, between *endogenous* and *exogenous stability*.[13] These two notions sort factors – as well as their origins – in two kinds: those that are internal to the schema or structure under investigation and those that are external to it.

3.2.1. The endogenous stability of styles

In what follows, we will discuss two aspects of the endogenous stability of styles, which is achieved through stabilization techniques. The first aspect is the coherence amongst statements. The second is the coherence between ideas, materialities, and institutions, all understood in a broad sense. To be sure, these two aspects do not refer to different things. On the contrary, the coherence of statements can be considered part of the coherence among ideas and other things.

3.2.1.1. Statements and (candidacies for) truths

In a previous subsection (3.1.1), we presented a fourfold structure that aspired to depict Hacking's use of history in his styles project. As we saw in the beginning of 3.1.2, the introduction of positivity is Hacking's second step, after his turn to Crombie, towards the substantiation of his philosophical project. His discussion of meaning and truth is indebted to two influences: analytical philosophy of language, on the one hand, and Comte and Foucault, on the other. What interests us here is that Hacking does not just use Foucault and Comte *simpliciter*. His approach to both, and especially to the former, is significantly mediated by analytic philosophy of language. At the same time, Hacking approaches philosophy of language in light of some Foucauldian insights. In other words, when Foucault talks about epistemes, he appears from Hacking's perspective to employ a coherentist theory of truth. More forcefully, his treatment of epistemes can be seen as a coherentist approach to truth substantiated by history. This is how Hacking uses Foucault's *Les mots et les choses* in his styles project. Thus, Foucault is employed as a historical philosopher of science and knowledge. In that sense, Hacking's discussion of truth is not just a philosophical idea; rather, it is a historiographical tool: it explains how

[12] See sections 3.2.2.1 and 3.2.2.2 below.

[13] This distinction is borrowed from Dries 2010.

scientific statements within a style hang together, and thus, to a certain extent, explains the stability of styles.

A style of reasoning is the condition of possibility of positivity: it puts forward "propositions that are up for grabs as true-or-false" (Hacking 2002b: 164). Furthermore, the meaning of these propositions is determined by the style in which they emerge (160). As such, a style determines both their truth conditions and the method for determining their truth value. In other words, a style determines the kinds of reasons we provide – along with the way we provide them – for justifying a proposition. Thus, the existence of these propositions, their meaning, and the method for verifying their truth are grounded in and depend upon a style of reasoning. In other words, meaning and method are contextually determined, and the context in this case is the style. It follows that the rationality of a style of reasoning depends on nothing else but on the style of reasoning itself; there are no external criteria to be invoked (167). This implies that the truth of style-dependent propositions is better described by a coherentist theory of truth (191). Thus, one of the techniques that make styles stabilize themselves is the formulation of statements that cohere together.[14]

3.2.1.2. Ideas, materialities, and institutions

As we have seen, with his turn to Crombie's notion of styles, Hacking refers to a specific "division of labor":

> Crombie's volumes [the *historian's* work] will, I hope, be read in part as an account of how conceptions of objective knowledge have come into being, while the *philosopher* can describe the techniques which become autonomous of their historical origins, and which enable styles of reasoning to persist at all. *Yet I would not push this division of labor too far.* (Hacking 2002c: 198-199, emphasis added)

The enterprise is philosophical, as the double question of the autonomy and persistence of stabilization techniques brings together three key issues of the philosophy of science: inevitabilism and contingency, realism and nominalism, and the explanation of the stability of scientific theories.

However, Hacking's approach to answering these questions is historical. He looks into Crombie's history and singles out the common elements of all styles that are deemed necessary for their stability. Furthermore, these elements become historiographical tools by means of which relevant data can be mined

[14] "The apparent circularity in the self-authenticating styles is to be welcomed. It helps explain why, although styles may evolve or be abandoned, they are curiously immune to anything akin to refutation. There is no higher standard to which they directly answer. […] [§] I believe that understanding the self-authenticating character of styles of reasoning is a step towards grasping the quasi-stability of science." (Hacking 2002c: 192).

out of the historical record. More importantly, these historiographical ideas guide Hacking himself to examine the historical record when discussing individual styles; and, moreover, this allows an interplay between the historical and the historiographical, or the philosophical and the historiographical. The hermeneutic aspect of Hacking's styles project is retained. This means that in light of new historical and historiographical material modifications can occur. For example, an element that is present in only one style can be dropped out as not being a common, constitutive element of stabilization; another can be introduced, as we come to realize that it names a group of elements common among different styles, and so on. The fact that Hacking presents only the outcome of his research should not obscure its dynamic character; on the contrary, we should keep in mind that that outcome is the result of a hermeneutic process. Finally, this also allows Hacking to make second-order, metahistorical remarks of the following sort: "Each style of reasoning has its own characteristic self-stabilizing techniques. […] [§] Almost the only thing that stabilizing techniques have in common is that they enable a self-authenticating style to persist, to endure" (*ibid.*: 193).

We can now turn to the very elements that Hacking refers to. Although he does not make this distinction, he presents *discursive* and *non-discursive* elements: he refers, on the one hand, to ideas, theories, and "marks (including data and data analysis)", and, on the other, to "material, institutional requirements" (*ibid.*: 194). His main idea can be understood as an extended version of the Duhem-Quine thesis.[15] Following Pickering, Hacking expands Duhem's confirmation holism to include both the set of non-discursive elements and the relation between discursive and non-discursive ones.[16] This expanded holism describes the structure of styles' self-stabilizing techniques; and it is the structure of what we called endogenous stability.

These two aspects of endogenous stability, that is, the coherence among statements and the extended confirmation holism, raise a philosophical issue with historiographical consequences. As we saw, Hacking claimed that although styles emerge from local, historical and social, contingent conditions, they are not reduced to them. The question that arises here is whether these self-stabilization techniques are sufficient for the autonomy of a style. More forcefully put, given those two endogenous stability aspects, the autonomy of a style is based on the autonomy of the corresponding intrastylistic reason. The question raised concerns the ontological status of this reason, and, specifically,

[15] Duhem is both a historical *philosopher* of science and a philosophical *historian* of science. His eponymous thesis appears in one of his primarily philosophical works (1914).

[16] Cf. Hacking 1992.

whether it is grounded into a metaphysical foundation, so as to transcend the contingent conditions from which it emerged. As we will see, the answer to this philosophical question goes hand in hand with a particular historiographical outlook.

3.2.2. The exogenous stability of styles

The explicitly acknowledged Whiggish character of Hacking's approach frames the discussion of exogenous stability (Hacking 2012: abstract, 599). First, it is not only his tool that is Whiggish. His whole approach is philosophical and as such, according to Hacking, it is Whiggish (602). Second, in his 1992 essay on styles Hacking explicitly associates his approach with a Foucauldian history of the present (Hacking 2002c: 182). However, elsewhere, in an essay not belonging to the styles project, he explicitly describes Foucault's history of the present in terms of contingency, and contradistinguishes it from Whiggism (Hacking 2002a: 24). The contingency indicated can be read as aiming to undo the metaphysical foundation on which inevitabilist narratives are grounded. Still, in the 1992 essay, a paragraph later, Hacking does refer to the presentist aspects of the project (Hacking 2002c: 183), while in his later assessment he takes explicitly some distance from Foucault,[17] a move that we should take at face value.

Finally, the inevitabilist aspects of Hacking's project have specific consequences concerning his use of history for philosophical purposes. The notion of inevitabilism is grounded in a series of metaphysically realist elements that provide the exogenous stability of a style. As we will see, these elements seem to belong either to a quasi-transcendental structure of the subject, or to an equally ahistorical world outside this subject. Hacking's reference to ecological history encapsulates both of these characteristics (Hacking 2012: 607). In what follows, we discuss three of these elements: (i) truth, (ii) objects, and (iii) human nature.

3.2.2.1. Truth and truthfulness

First, truth. According to Hacking, Bernard Williams' *Truth and Truthfulness* "suggests a way to explicate the autonomy of Crombie's styles, in a way that is aligned to discussions of truth by analytical philosophers" (*ibid.*: 605). Hacking parallels his own distinction between truth and being-a-candidate-for-truth with Williams's between truth and truthfulness. Williams's referring to a quadruple relation among "truth, language, meaning and belief" describes

[17] "*One could cloak the styles project in the mantle of a recently trendy phrase used for a short time by Michel Foucault, 'history of the present'. That would be pretentious: the shoe does fit, but it is for dancing only.* A self-conscious use of the past to reflect on the present has all sorts of dangers, but philosophers are in the business of living dangerously." (Hacking 2012: 602, emphasis added).

an ahistorical schema. The content of each category might vary historically – and both Hacking and Williams admit that it actually does – , yet the very interrelation among these categories is fixed and invariable (*ibid.*).

However, this parallel is neither uncontroversial nor innocuous. Hacking, following Williams, admits the existence of an ahistorical structure underpinning styles, which are otherwise historically constituted. Thus, Hacking allows for an ahistorical element to guarantee the stability of a style. In other words, the autonomy of styles is explained in virtue of something external to them.

Hacking, again following Williams, admits that truth cannot be identified with justification, and this idea is underpinned by a metaphysical, extrahistorical commitment. Thus, there is a tension between Hacking's reassurances that styles "do not answer to some other, higher, or deeper, standard of truth and reason than their own [...] to some external canon of truth independent of itself," and the extrahistorical structure of "truth, language, meaning and belief" (*ibid.*). In other words, there is a tension between Hacking's reassurances that crystallization and sedimentation are the reasons for the autonomy of a style and this extrahistorical commitment.

Again, we do not want to score philosophical points. This tension is important for our overall point. Hacking's invoking of Williams has a problematic aspect. Williams's distinctions seem to be used less for the mustering of new material, than for underpinning the stabilization techniques presented. Moreover, although Hacking does not suggest an explicit causal link between this ahistorical structure of truth and the ecological and cognitive structures he refers to, the former can be interpreted as being grounded in the latter. The reference to these aspects is important, as Hacking's historical philosophy of science seems to acquire at the same time a realist, inevitabilist, and internalist orientation.

3.2.2.2. Objects

According to Hacking, "[e]arly on, the styles project maintained that each style of scientific thinking & doing introduces a new class of *objects*" (*ibid.*: 606, emphasis added). Furthermore, he maintained that the realism-antirealism questions, at least concerning some entities, are style specific (*ibid.*). That is, questions concerning the ontological status of entities are a byproduct of the styles within which these entities have been proposed. Hacking, however, strongly qualifies his view: "This *does not mean that objects of the class did not exist* before there was a way to investigate them. *That is nonsense.* Each new style of thinking & doing introduced a new class of objects into discourse" (*ibid.*: 606, emphasis added). Styles introduce, *mainly*, classes of objects and not objects themselves.

Second, as we saw, Hacking does not suggest that styles are a prerequisite for having knowledge. Styles produce knowledge *par excellence*, yet there is knowledge outside styles. Hacking accepts the existence of style-independent observational statements. Traditionally, these are considered to have a relation of correspondence with the facts they describe, and, in this sense, he allows for a different kind of theory of truth and meaning from the one applied in the case of style-dependent propositions.

Third, there is another element that links these two ideas – that is, the ontological independence of the objects of a class introduced by a style, and extrastylistic propositional knowledge. Hacking allows the possibility of referring to extrastylistic objects in the context of explaining the problems encountered within a style; and this possibility may substantiate his earlier acknowledgement of the need for external input to explain the demise of a style. In other words, extrastylistic objects and extrastylistic knowledge about them may constitute the external input in question.[18]

Hacking's points raise the following philosophical questions: Is such a clean-cut distinction between style-independent and style-dependent knowledge possible at all? Even if the answer is a positive one, what is the relation between these two kinds of knowledge? Moreover, if objects exist independently of their classes, don't they provide constraints to possible classifications? These questions are interrelated, and their answer harks back to a more fundamental one: the question of the perspective from which Hacking talks. He needs a metaphysical standpoint from which these points can be made. Again, these points do have historiographical consequences. Hacking's acknowledged invoking of external input involves an internalist perspective, while it compromises the contingency he attributes to styles, allowing for inevitabilist accounts.[19]

3.2.2.3. Ecological history, cognitive history, and philosophical anthropology

Hacking raises the question of a kind of ultimate explanation, of the "larger grounds", of the more fundamental conditions of possibility for the presence, stabilization and persistence of styles (Hacking 2012: 600). Moreover, we can trace Hacking's endorsement of an almost ahistorical conception of human nature. According to his earlier 1992 account, these conditions are the subject matter of philosophical anthropology. However, referring mainly to the later Wittgenstein, Hacking characterizes these conditions as "brute". They describe a stock of general platitudes about "human beings and their

[18] This idea is traced in Hacking's critical insight concerning Foucault's epistemes that we discussed in section 3.2 above.

[19] As earlier, the notion of internalism refers to a mode of explanation that does not invoke social or other elements considered external to scientific inquiry.

place in nature", and they are extrastylistic and extrascientific (Hacking 2002c: 196-197). As there is little we can say about them, for earlier Hacking philosophical anthropology is rather thin.

In his later 2012 account, these conditions are spelled out in terms of ecological factors. In the beginning of his paper, in order to retain his middle position on the contingency of styles, Hacking makes an explicit distinction between man's psychophysical setup and the local context within which this setup is at play. The latter is the purely contingent factor, while the former is a more universal one, albeit biologically contingent. However, Hacking's rhetoric mitigates this contingency. The "specific local settings are" described as "*grounded in human capacities* that are presumed to be *universal*" (Hacking 2012: 600, emphasis added). In short, an ultimate, ecological explanation of the presence (*qua* emergence, stabilization, and persistence) of styles is grounded in human nature.

Later in the same paper, this idea becomes more explicit. Philosophical anthropology remains extrastylistic, but now the historian can flesh out the content of human nature. The analysis of the conditions of possibility of styles now becomes the task of the historical philosopher of science, if not of the historian of science (*ibid.*: 608). The earlier platitudinous description of "human beings and their place in nature" becomes now the austere set of "biologically cognitive facts" (607). Furthermore, this cognitive setup constitutes the bedrock in which certain structures within styles are grounded (607). However, this setup should not be conceived as a static structure; it should be understood as the outcome of an evolutionary process and of its interrelation with both natural and technical environment (607, §20).

Thus, Hacking invokes a series of authors – Scott Atran, Philippe Descola and Pierre Hadot – and makes use of their philosophical ideas – "biologically cognitive facts", "innate module[s]", "a long view of the idea of nature" (*ibid.*: 607 and 608), respectively – as metaphysical foundations for the explanation of the stability of styles.

This imagery raises significant philosophical issues that in turn have specific historiographical consequences. First, as we saw at the end of the previous subsection (3.2.1), the question arises whether the self-authentication techniques associated with a style are alone sufficient to justify its independence from the social conditions out of which it emerged. It seems now that Hacking answers this question in the negative: an external aspect is needed – in this case a revamped version of the human mind – to ground the autonomy of intrastylistic reason. In other words, Hacking claims that styles emerge from "local microsocial incidents", yet they are not reduced to them. The autonomy of intrastylistic reason cannot be guaranteed by its self-stabilization

techniques alone. It is predicated upon the existence of universal "human capacities" in which these techniques are grounded.

Hacking's middle position between contingency and inevitabilism is compromised in favor of a more inevitabilist stance. To be sure, human capacities are contingent. Evolutionary development does not entail any predetermination. These very capacities, however, and the foundationalist role that Hacking ascribes to them close down a range of possibilities. The need of invoking extrastylistic factors for the explanation of the development and demise of styles, and the idea that the autonomy of stylistic reason is, ultimately, metaphysically grounded indicate that Hacking is more orientated towards the inevitabilism of the "foreordained right route" he wants to avoid (*ibid.*: 600).

4. *Concluding remarks*

In this paper we have undertaken a critical appraisal of the relation between history of science and philosophy of science in Ian Hacking's styles of scientific reasoning project. Specifically, we argued for three interrelated points. First, we distinguished between historical *philosophy* of science and philosophical *history* of science, and argued that Hacking's project belongs to the former. Second, we discussed the *confrontation* and *hermeneutic* models of HPS, and provided substantive evidence that in his styles project Hacking adopts a version of the latter. Furthermore, we showed that he adopts a strongly integrated approach to HPS, acknowledging, though, the autonomy of the historical material and the independence of historical facts from his philosophical perspective. Finally, we discussed the historiographical implications of his approach. Hacking aims at maintaining a middle position, on the one hand, between contingency and inevitabilism, and, on the other, between internalism and externalism in the explanation of the stability of scientific knowledge. We argued, however, that his philosophical and metahistorical commitments compromise his position towards a more inevitabilist and internalist orientation.

Acknowledgments

We are grateful to Matteo Vagelli's for the invitation to contribute to this special issue, as well as for his constructive comments on an earlier version of this paper. For years now, Ian Hacking's work has been a constant source of inspiration and enlightenment; none of the above reflections would have been possible without him. Manolis Simos is indebted to Nicholas Vrousalis and Manuel Dries for helpful and insightful discussions on several of the

issues examined in this paper. Finally, his research for this paper was co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Reinforcement of Postdoctoral Researchers – 2nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (ΙΚΥ).

Manolis Simos
Department of History and Philosophy of Science,
National and Kapodistrian University of Athens
esimos@phs.uoa.gr

Theodore Arabatzis
Department of History and Philosophy of Science,
National and Kapodistrian University of Athens
tarabatz@phs.uoa.gr

References

Arabatzis, Theodore, 2006, "On the inextricability of the context of discovery and the context of justification", in J. Schickore and F. Steinle, eds., *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*, Archimedes 14, Springer, Dordrecht: 215-230.

Arabatzis, Theodore, 2012, "Hidden entities and experimental practice: Renewing the dialogue between history and philosophy of science", in S. Mauskopf and T.M. Schmaltz, eds., *Integrating History and Philosophy of Science: Problems and Prospects*, Springer, Dordrecht: 125-139.

Arabatzis, Theodore, 2017, "What's in it for the historian of science? Reflections on the value of philosophy of science for history of science", in *International Studies in the Philosophy of Science*, 31, 1: 69-82.

Dries, Manuel, 2010, "On the logic of values", in *Journal of Nietzsche Studies*, 39: 30-50.

Duhem, Pierre, 1914, *La théorie physique son objet et sa structure*, Chevalier et Rivière, Paris; Eng. tr. by Philip P. Wiener 1954, *The Aim and Structure of Physical Theory*, Princeton University Press, Princeton.

Hacking, Ian, 1992, "The self-vindication of the laboratory sciences", in A. Pickering, ed., *Science as Practice and Culture*, The University of Chicago Press, Chicago: 29-64.

Hacking, Ian, 2002, *Historical Ontology*, Harvard University Press, Cambridge, Massachusetts.

Hacking, Ian, 2002a, "Historical ontology", in Hacking 2002: 1-26.

Hacking, Ian, 2002b, "Language, truth, and reason", in Hacking 2002: 159-177.

Hacking, Ian, 2002c, "'Style' for historians and philosophers", in Hacking 2002: 178-199.

Hacking, Ian, 2012, "'Language, truth, and reason' 30 years later", in *Studies in History and Philosophy of Science*, 43, 4: 599-609.

Kindi, Vasso, 2014, "Taking a look at history", in *Journal of the Philosophy of History*, 8, 1: 96-117.

Lakatos, Imre, 1980, "History of science and its rational reconstructions", in J. Worrall and G. Currie, eds., *The Methodology of Scientific Research Programmes: Philosophical Papers, Vol. 1*, Cambridge University Press, Cambridge: 102-138.

Latour, Bruno and Woolgar, Steve, 1986, *Laboratory Life: The Construction of Scientific Facts*, Princeton University Press, Princeton.

Schickore, Jutta, 2011, "More thoughts on HPS: Another 20 years later", in *Perspectives on Science*, 19, 4: 453-481.

Vagelli, Matteo, 2019, "Historical epistemology and the 'marriage' between history and philosophy of science", in E. Herring *et al.*, eds., *The Past, Present, and Future of Integrated History and Philosophy of Science*, Routledge, London: 96-112.

Were experiments ever neglected? Ian Hacking and the history of philosophy of experiment[1]

Massimiliano Simons and Matteo Vagelli

Abstract: Ian Hacking's *Representing and Intervening* is often credited as being one of the first works to focus on the role of experimentation in philosophy of science, catalyzing a movement which is sometimes called the "philosophy of experiment" or "new experimentalism". In the 1980s, a number of other movements and scholars also began focusing on the role of experimentation and instruments in science. Philosophical study of experimentation has thus seemed to be an invention of the 1980s whose central figure is Hacking. This article aims to assess this historical claim, made by Hacking himself as well as others. It does so first by highlighting how a broader perspective on the history of philosophy reveals this invention narrative to be incorrect, since experimentation was a topic of interest for earlier philosophers. Secondly, the article evaluates a revision of this historical claim also made by some philosophers of experiment: the rediscovery narrative, which frames Hacking and others as having rediscovered the work of these earlier authors. This second narratives faces problems as well. Therefore we develop a third narrative which we call the contextualist narrative. Rather than considering experimentation in an essentialist manner as a fixed research object that is either present or not in the work of specific authors, experimentation should be addressed through a narrative that asks in what way it becomes a philosophical problem for certain authors and for what purpose. Such contextualization enables a repositioning of Hacking's philosophy of experiment in relation to the specific debates in which he intervened, such as the realism-antirealism debate, the Science Wars and the debate on incommensurability.

Keywords: Ian Hacking; philosophy of experiment; experimentalism; realism; constructivism; Science Wars

1. *Introduction*

Ian Hacking's work is often credited as being one of the main inspirations of the "philosophy of experiment" (Hacking 1988a) or "new experimentalism"

[1] The authors are named alphabetically, since the work on the paper was shared equally. Simons wrote pages 170-177 and Vagelli wrote pages 177-186. The authors want to thank the organisers and participants of two conferences where earlier versions of this paper were presented: "Open Epistemologies: Mach, Bachelard, Feyerabend (Lisbon, 20-21 September 2019)" and "Bachelardismes et antibachelardismes en France: Controverses épistémologiques des années 1960 (Paris, 16-17 April 2019).

ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.339

(Ackermann 1989), a philosophical program that has put the role of experiments in science on the philosophical agenda. That this movement has gained influence and status over the past 40 years is evident in the numerous books that have appeared on the topic (Gooding and Pinch 1989; Mayo 1996; Radder 2003) and the birth of related philosophical subtopics, such as exploratory experimentation (Steinle 2002; Burian 2007) or simulation (Lenhard 2007; Winsberg 2009).

Often Hacking's *Representing and Intervening* (1983) is taken as the starting point for this movement. Allan Franklin, a pioneering representative of the new experimentalists, maintained that "*Representing and Intervening* has made it legitimate to discuss the philosophy of experiment" (Franklin 1986: x). More recently, Theodore Arabatzis has stated that "it was only recently, during the 1980s, that experimental practice attracted the attention of philosophers of science" and that "Ian Hacking's work has been decisive in redressing the neglect of experiment and in bringing out its philosophical significance" (Arabatzis 2008: 162).

This historical claim can be found in Hacking's book itself. In the second part of *Representing and Intervening*, Hacking stresses that science not only represents the world but also intervenes in it. Hacking links this with the historical claim that the topic of experimentation has been ignored in the philosophy of science. "Philosophers of science constantly discuss theories and representation of reality, but say almost nothing about experiment, technology, or the use of knowledge to alter the world" (Hacking 1983: 149).

This historical claim also has been defended by other new experimentalists. For instance, in his *The Neglect of Experiment,* Franklin asks "Who was neglecting experiment? Certainly not scientists. I believed then that it was historians, philosophers, and sociologists of science [….] Actual experiments were rarely discussed" (Franklin 1986: 1). Similarly, Robert Ackermann starts his analysis by saying that "[p]revious views have left the role of scientific experimentation out of account" (Ackermann 1985: 30).

What is perhaps more noteworthy is that new experimentalists were not the first to make this historical claim about a gap in knowledge regarding experimentation. The late 1970s marked the beginning of a period characterized by what is often called a "practical" or "practice turn" (Schatzki, Knorr-Cetina and von Savigny 2001, Soler *et al.* 2014, Agazzi and Heinzmann 2015), witnessed by a sudden rise in interest in the experimental, technical, and material aspects of science.

Early influential examples of this turn in the sociology and history of science include Andrew Pickering's *Constructing Quarks* (1984), Steven Shapin's and Simon Schaffer's *Leviathan and the Air-Pump* (1985), and Harry Collins'

Changing Order (1985). Pickering, for instance, aimed to introduce the agency of scientists into sociological accounts: "One gets little feeling that scientists actually *do* anything in their day-to-day practice" (Pickering 1984: 8). Similarly, Shapin and Schaffer opened their book with the statement that "[o]ur subject is experiment. We want to understand the nature and status of experimental practices and their intellectual products" (Shapin and Schaffer 1985: 3).

Around the same time, sociologists began performing ethnographic studies of laboratories, again stressing the role of experimentation and intervention. The most famous example is *Laboratory Life* (1979) by Bruno Latour and Steve Woolgar, but others soon followed (Knorr-Cetina 1981; Lynch 1985; Traweek 1988). Like the new experimentalists, they stressed the innovative and revolutionary nature of their work and the goal of correcting the far-too-theoretical views dominating philosophy of science:

> What makes laboratory theories so atheoretical is the lack of any divorce from instrumental manipulation. Instead, they confront us as discursively crystallised experimental operations, and are in turn woven into the process of performing experimentation. (Knorr-Cetina 1981: 4)

Indeed, what most of these scholars and movements share is a form of self-description that stresses the *novelty* and *innovativeness* of their own approach: they advance the claim that experimentation, intervention, and instrumentation were not on the philosophical radar before they introduced it. The first section of this paper aims to evaluate this *invention narrative* and its relation to Ian Hacking.

In the second section, we show how problematic this invention narrative is from a historical point of view. We therefore go on to explore a second hypothesis: that the rise of a philosophy of experiment in the 1980s was due less to the *invention* than to the *rediscovery* of the forgotten and neglected subject of experimentation. As we will highlight, this *rediscovery narrative* also emerges, though less explicitly, in the work of the practical-turn protagonists cited in the invention narrative accounts above. To illustrate, we will focus on the example of Gaston Bachelard, exploring how his work on experimentation was taken up by more recent authors such as Hacking and Latour.

The example of Bachelard will also demonstrate, however, that the rediscovery narrative too faces problems. We therefore will argue for a third possibility, namely that it is better to abandon the idea that 'experimentation' has a fixed essence with the same significance for different periods of the history of the philosophy of science. We maintain that experimentation instead should be seen as an ambiguous, contextually informed resource that can be mobilized for multiple purposes.

Thus, we end by proposing a third account that eschews taking up experimentation as a fixed research object that is either present or not and instead offers a *contextualist narrative* centered around two questions: *in what way* did experimentation become a philosophical problem for certain authors and *for what purpose*? From this perspective, we suggest that what makes Hacking's claims especially innovative is how they reconceptualize a number of existing debates – such as those on realism vs. antirealism, the Science Wars, and incommensurability – by mobilizing experimentation as a resource.

2. *The invention narrative*

As remarked in the introduction, the emphasis on the experimental aspects of science that characterized new experimentalism was in many cases accompanied by an historiographical claim about the novelty of experimentation as a philosophical subject. This *invention narrative* is widespread among philosophers, historians, and sociologists of science. In this section, we will focus on two key versions of it, advanced by Ian Hacking and Peter Galison respectively.

2.1. Hacking's back-to-bacon movement

In a paper which anticipates many of the claims of *Representing and Intervening*, Hacking argues that

> no field in the philosophy of science is more systematically neglected than experiment. Our grad school teachers may have told us that scientific method is experimental method but histories of science have become histories of theory. Experiments, the philosophers say, are of value only when they test theory. Experimental work, they imply, has no life of its own. (Hacking 1982: 71)

A few years later, in a review paper, he expands on this narrative. According to Hacking, before the 1980s "there was almost no reflective philosophy of experiment", since philosophers and historians of science had "neglected the experimental side of science" (Hacking 1988a: 147). He adds that "what little had been published was not seen as writing about experiment – that was not something to write about – but as discussion of the theory/observation distinction, or the impossibility of eliminating a theory by crucial experiment, etc." (Hacking 1988a: 147). In later publications and interviews, Hacking does not hesitate to ascribe himself the role of trailblazer with respect to philosophical studies of experimentation:

> Learning from other people, I started the enthusiasm for experiment in the philosophy of science. My friend Francis Everett and I used to go walking in the Stanford

hills [...] He's the person who's planning the only experimental test of the theory of gravity. It's called Gravity Probe II [...] He and I started talking about experiments. It happened that *Representing and Intervening* came out just a little bit before everybody else's books on experiments except for Bruno Latour's *Laboratory Life*. (Hacking 1992a: 5)[2]

Thus in *Representing and Intervening*, Hacking hoped "to initiate a Back-to-Bacon movement, in which we attend more seriously to experimental science" (Hacking 1983: 150). But what he actually meant was something more than a simple return to Bacon's philosophy: Hacking's aim was to reaffirm the role of scientific experiments against the exclusive attention philosophers gave the nature, structure, functions, and limits of scientific *theories*. What Hacking further argues is that new experimentalism should not only account for experiments in science but moreover should assign them a more primary role with respect to theory. The controversial claim Hacking puts forward in this respect is that "experimentation has a life of its own" (Hacking 1983: 150), largely independent of the theoretical frameworks in which it occurs. Thus, Hacking's claim to novelty lies in his point that while earlier philosophers of course had already treated the topic of experiments, they always did so in relation to (or rather, in a manner dependent on) theory. Theory-independent experimentation was never considered in philosophical discussions, according to Hacking, and it is precisely this type of experimentation that he aimed to move to the spotlight. His appeal reminds us that if we cannot conceive of experimental practices in themselves, *qua* practices, and not as the expression, extension, or confirmation of some theory, we continue to miss a fundamental trait of scientific inquiry.

2.2. Galison's critical postmodern model

Next to Hacking's, the most influential instance of a new experimentalist invention narrative is the one put forward by Peter Galison. Galison's *How experiments end* (1987) was quickly recognized as a paradigmatic study of the new experimentalism. Like Hacking, Galison makes historical claims about when experimentation became a topic of concern for philosophers and historians. He begins the preface of the book with the following claim: "Despite the slogan that science advances through experiments, virtually the entire literature of the history of science concerns theory" (Galison 1987: ix).

Galison elaborated on this history more fully a year later by comparing three philosophical models of the nature of scientific inquiry: the positivist, anti-pos-

[2] In other places, making reference to Ravetz (1971), he seems to grant that sociologists were the real initiators of the study of experimentation: "once people did begin to think about experiment, those conducting social studies of science got there first" (Hacking 1988: 148).

itivist, and 'critical postmodern' (Galison 1988). For Galison, positivists begin with a dual-layered image of science, with a shifting lower layer of theories and a steadily growing higher layer of observations, which shapes these theories and their evolution. The anti-positivists later inverted that picture, instead conceiving of observation as determined by theory and therefore arguing that, like theories, a set of observations could shift in light of a new theory.

Galison disagrees with both models and links new experimentalism with a wholly pluralistic historiographic view – which he calls the critical postmodern – characterized by a plurality of levels corresponding to globally cooperating but also partially autonomous layers. Not only does Galison disagree with the assumption of the unity of science at work in both of the previous models, he also breaks down the dichotomy between theory and observation, substituting the latter for two new layers: experiments and instruments. Experimentation is thus explicitly discussed, by Galison, in relation to the material culture of instruments and experimental materials.

3. *The rediscovery narrative*

In the previous section we outlined two examples of new experimentalist arguments claiming to offer something new in philosophy of science, namely the foregrounding of the role of experiments and instruments in science. Yet though it rarely has been questioned (for an exception, see Radder 2009), this claim can be problematized historically. Indeed, it is far from true that experiments and instruments were never a proper object of philosophical concern before the 1970s.

There are many examples of philosophical engagement with experimentation dating to the end of the 19th century or first decades of the 20th. One thinks of Ernst Mach or Ludwik Fleck in Germany, Pierre Duhem, Henri Bergson, or Gaston Bachelard in France, and John Dewey or P. W. Bridgman in the Anglo-American context. One example worth highlighting is Hugo Dingler and his book *Das Experiment: Sein Wesen und seine Geschichte*, in which we find an early example of the invention narrative, predating the above examples by fifty years: "A real 'philosophy of experiment' has never been written to my knowledge. Therefore this book should at the same time be seen as a pioneering study in this area" (Dingler 1928, i).

We will only briefly explore one such case, that of Gaston Bachelard, because Bachelard's philosophy of experiment was explicitly taken up by Hacking and others (3.1). This example underscores the invention narrative as problematic and points toward a plausible alternative hypothesis, which we call the *rediscovery narrative*. However, as we will see, this alternative hypothesis also

faces problems, most notably in its inability to account for the varying ways in which Bachelard's work has been interpreted by later philosophers or has been mobilized to support very different, even opposing, claims (3.2).

3.1. The eternal return of experiments: the case of Gaston Bachelard

Bachelard's philosophy of science included a "philosophy of experiment" (Bachelard 1927: 26), which he mainly conceptualized through the concept of *phénoménotechnique*, first introduced in the early 1930s. According to Bachelard, contemporary science was characterized by a shift away from purely descriptive phenomenology toward "a phenomenotechnique through which new phenomena are not simply found but invented, constructed and built from all parts" (Bachelard 1931: 76). His philosophy thus defended a form of constructivism, in the sense it maintained that "science does not correspond to a world to be described. It corresponds to a world to be constructed" (Bachelard 1951: 46). One of the examples Bachelard provides in support of this claim is that of isotopes in mass spectroscopy:

> The trajectories that allow the separation of isotopes in the mass spectroscope do not exist in nature; one must produce them technically. They are reified theorems. We shall have to show that that which man makes by a scientific technique […] does not exist in nature and neither does a natural range of natural phenomena. (Bachelard 1949: 103)

In Bachelard's work, we thus find a clear philosophy of experiment. Although authors such as Althusser, Foucault, and Bourdieu – who mainly mobilized Bachelard's idea of an epistemological rupture (i.e. of a radical break between scientific thinking and imagination; see Simons 2018) – largely ignored this part of Bachelard's philosophy, other prominent philosophers did take it up. This is the case, for instance, of Georges Canguilhem (1955), Gilbert Simondon (see Bontems 2010), and François Dagognet (1965; 1979), who, in their respective analyses of different sciences drew from Bachelard's notion of phenomenotechnique. Their attention to Bachelard's work undermines the plausibility of a "rediscovery narrative", since Bachelard's philosophy of experiment was never forgotten.

However, it is also true that from the 1980s on there has been growing interest in Bachelard's phenomenotechnique, prevalently within Anglophone philosophy of science (Tiles 1984, Castelao-Lawless 1995, Chimisso 2001, Simons 2018). Hans Radder has even framed one of the central issues in the philosophy of experiment as the 'Bachelardian challenge': "it is the question how scientific knowledge can be about a human-independent reality, if this reality is so thoroughly dependent on human work" (Radder 1993: 328). In a similar vein,

Hans-Jörg Rheinberger has picked up the notion of phenomenotechnique to articulate the experimental aspects of molecular biology (Rheinberger 2005).

In new experimentalism, we also find elements of this rediscovery of Bachelard's phenomenotechnique. Galison makes occasional references to Bachelard; he describes the latter as "a gently materialist opponent of a certain stripe of neo-Kantian idealism" (Galison 1997: 18n24). Similarly, Ackermann uses Bachelard's work to explore the idea that scientific observations are discontinuous with common sense (Ackermann 1985: 88).

There is no mention of Bachelard in Hacking's *Representing and Intervening.* To some extent this is to be expected: Hacking's interest in shedding light on the power of experiments outside the limits of theory would find little of use in Bachelardian ideas of instruments or of theoretical entities as "reified theories". However, in a text whose first appearance dates to 1983, Hacking acknowledges that his position is "strikingly similar" to the one advanced by Bachelard's *Le matérialisme rationnel* (Bachelard 1953; see also Vagelli 2014: 262). There, Hacking notes, Bachelard pointed out that the introduction of new phenomena, such as the photoelectric effect, represented an "absolute discontinuity" with the history of science. However, as Hacking further argues, Bachelard also believed in scientific progress and in the accumulation of experimental techniques (Hacking 2002: 45). This latter point aligns with Hacking's idea that the phenomena produced in a scientific laboratory have the ability to persist regardless of changes in theory (Hacking 1983: 220-233). Scientific effects are relative in the sense that they are bound to our ability to recreate them; they depend on our technical skills but still cut across different theories and styles of reasoning. Phenomena that are stable, that is, that are capable of being regularly reproduced in a laboratory setting, are not objective in the absolute or foundationalist sense, because they still depend on our technical skills and on the invention of our experimental apparatuses, but they are objective in the sense that they can be largely independent of general theories.

Thus, if we consider the example of Bachelard, a *rediscovery narrative* does seems a more accurate assessment of Hacking's contribution to the philosophy of experiment than the *invention narrative*: experiments were forgotten, and new experimentalists put the topic back on the table. And indeed, we find this rediscovery narrative in several overviews of the philosophy of experiment. Friedrich Steinle for example, starts an article by discussing how Francis Bacon, John Stuart Mill, and even Pierre Duhem vividly discussed the topic of experimentation before it fell off the philosophical radar:

> Throughout the 20th century, however, philosophy of science narrowed its perspective on experiment significantly. […] Only in the 1980s, did philosophy of science

again take up the question of experiment. The 'New Experimentalism' in philosophy of science arose, stimulated by Hacking's emphasis on a 'Baconian variety' of experiment, clearly emphasizing the insufficiency of the older accounts. (Steinle 2002: 408-409; cf. Feest and Steinle 2014: 274)

Nevertheless, this rediscovery narrative also faces problems. The work of earlier authors is often not addressed or analyzed in a uniform manner, leading to very divergent interpretations of earlier thinkers. We will illustrate this weakness by contrasting Hacking's interpretation of Bachelard with that of Bruno Latour. The comparison will show that it remains unclear exactly *what* is being rediscovered in the first place.

3.2. Divergent interpretations of Bachelard

The problem with the rediscovery narrative is its assumption that there is a fixed object (the 'experiment') that can be rediscovered. We want to question this assumption, again using the example of Bachelard. Although numerous authors have taken up the Bachelardian notion of phenomenotechnique, they often interpret it in radically different ways. As we saw above, Hacking, for example, advanced a relatively realist interpretation of Bachelard's philosophy (see Vagelli 2017). In the work of someone such as Bruno Latour, however, we find a quite different Bachelard.

Latour and Woolgar (1979) took inspiration from Bachelard's notion of phenomenotechnique to support the claim that facts are artificial in the sense that they are manufactured (as opposed to phony). As has been well-noted, Latour and Woolgar made this point using the example of the laboratory synthesis of TRF (Thyrotropin Releasing Factor), a paradigmatic case of the social construction of a scientific fact.

We may think TRF has been there all along, just waiting to be discovered, but Latour and Woolgar argue that it is only after 1969 and a particular series of laboratory events, exchanges, and negotiations that it became a fact. At this historical juncture, scientists decided to turn a statement about the chemical structure of TRF into an object, which then came to be seen as the cause of the statement. Yet since scientific knowledge is sustained by the network of creators and distributors of that knowledge, a change in the context of laboratory norms might turn TRF back into an artefact (or a sentence). It is in this context that Latour and Woolgar take up Bachelard's notion of phenomenotechnique:

> It is not simply that phenomena depend on certain material instrumentation; rather, the phenomena are thoroughly constituted by the material setting of the laboratory.

The artificial reality, which participants describe in terms of an objective entity, has in fact been constructed by the use of inscription devices. Such a reality, which Bachelard (1953) terms the 'phenomenotechnique,' takes on the appearance of a phenomenon by virtue of its construction through material techniques. (Latour and Woolgar 1979: 64)

The contrast between this seemingly constructivist stance and Hacking's realist position is striking, and it is confirmed further if we analyze the distinction Hacking draws between scientific facts that are historically *contingent* and those that are historically *constituted*. Hacking's realism is built on the belief that phenomena are created – that is, they are "brought into being at moments of time" – but that they cannot be said to be historically constituted, because they existed before becoming objects of scientific inquiry and "are phenomena thereafter, regardless of what happens" (Hacking 2002: 44). This idea too is supported by some of Bachelard's texts, for instance when Bachelard writes:

The electron existed before the 20th century men and women. But before them, the electron did not sing. In the triode valve, however, the electron sings. This phenomenological realization occurred at a precise point when mathematical and technical development was coming to maturity. (Bachelard 1938: 246)

With both Hacking and Latour and Woolgar supporting their claims by drawing from Bachelard's work, we end up with two competing interpretations of phenomenotechnique. One way out of this conflict would be to investigate which interpretation of Bachelard is correct and then subsequently to assess whether Hacking's philosophy of experiment was a faithful rediscovery of Bachelard's earlier work or a (perhaps fruitful) misinterpretation of it. We would like to draw a different lesson from this debate, however: namely, that it is worth questioning the assumption that there is a fixed philosophical notion called 'experiment' to be rediscovered through the work of Bachelard.

4. *The contextualist narrative*

One possible rebuttal of the previous arguments is to say that, though there have been earlier philosophers of experiment who were never forgotten, the value of Hacking's new experimentalism resides in the rediscovery of this tradition in the Anglo-American context. Although this counterclaim is partly true, it at best concludes that new experimentalism was nothing novel in the history of the philosophy of experiment but rather a mere product of the 'globalized parochialism' of Anglophone philosophy and its impulse to ignore other traditions (Wolters 2015).

Another possible counterargument would be to say that a philosophy of experiment was present in Bachelard's work in only an embryonic state – and that only in the 1980s was a *proper* philosophy of experiment developed. Such an assessment finds grounds in Latour's claim that, although a step in the right direction, Bachelard's "interest in demonstrating the 'mediations' in scientific work was never extended" (Latour and Woolgar 1979: 258). However, this line of argument risks anachronism, as it assesses Bachelard's philosophical project according to the standards of a "good" philosophy of experiment as they developed from the 1980s on. Moreover, this essentialist view of what a philosophy of experiment should look like also does injustice to Hacking's originality. The value of Hacking's philosophy of experiment dwells in something other than filling in the details of Bachelard's earlier work. Indeed, we argue that it resides above all in the innovative ways Hacking mobilizes the topic of experimentation to introduce new philosophical options into well-entrenched debates. We therefore propose a *contextualist narrative* of the history of the philosophy of experiment, which we think is able to solve the problems posed by the earlier narratives while simultaneously acknowledging Hacking's unique contribution.

4.1. Resituating Hacking's work on experiments

Several scholars have noted that Hacking's famous slogan that "experimentation has a life of its own" can mean various things (Mayo 1996: 62; Soler *et al.* 2014: 7-8). Typically, it is associated with a critique of the theory-centeredness of philosophy of science: the purpose of experimentation often diverges from the mere testing of a general theory, often instead consisting in the aim of better articulating phenomena or of simply making certain that instruments work. We do not intend to contest this claim about the criticism of theory-centeredness, but by advancing a contextualist narrative we mean stress that this interpretation of Hacking leaves open the question of *why* these aspects of experimentation are philosophically relevant. It is here that context matters, since new experimentalists find it important to invoke these other non-theory-oriented dimensions of experimentation in response to specific problems found in the philosophy, history, and sociology of science.

In the framework of the sociology of scientific knowledge, authors such as David Bloor and Barry Barnes, and later Collins, Pickering, and Shapin, argued that scientific controversies are never determined by purely logical or rational means, but rather that social factors play a decisive role. In their *Leviathan and the Air-Pump*, for instance, Shapin and Schaffer argue that the production of accepted matters of fact "rested upon the acceptance of certain social and discursive conventions, and that it depended upon the production and protection of a special form of social organization" (Shapin and Schaffer 1985: 22). In

relation to experimentation, this idea was expressed most famously by Collins in his notion of 'experimenters' regress':

> The problem is that, since experimentation is a matter of skillful practice, it can never be clear whether a second experiment has been done sufficiently well to count as check on the results of a first. Some further test is needed to test the quality of the experiment – and so forth. (Collins 1985: 2)

According to Collins, whether or not an experiment is accepted as correct is, in the end, based on social consensus rather than rational argument. To make such claims, these sociologists relied on philosophical arguments about underdetermination or the empirical equivalence of different theories, taking inspiration from scholars such as W.V.O. Quine and Nelson Goodman (see Zammito 2004). It is in this sense that theory-ladenness becomes a problem: if the correctness of a theory can never be determined based on strict deductive or empirical arguments, the door remains open for sociological explanations.

It is in the context of this discussion that we can identify the value of new experimentalism as "having provided us sticks with which to beat the social constructivists" (Mayo 1996: 61). In response to the sociologists, the new experimentalists aimed to restore the constraining role of empirical evidence but in a *novel* way, moving away from traditional philosophers of science who accepted a theory-centered model of science toward the rabbit hole of underdetermination. This is indeed the type of self-positioning and framing one finds in the prefaces of Ackermann (1985), Franklin (1986), and Galison (1987). The strategy of these authors was to stress that there are no given, absolute logical and empirical constraints but that they can be *introduced* – and this is precisely what *intervening* in science is all about.

Galison, for instance, states that "there is no *strictly logical* termination point inherent in the experimental process" nor is there "a universal formula for discovery, or an after-the-fact reconstruction based on an inductive logic" (Galison 1987: 3). Nonetheless, his conclusion is not that social factors therefore determine outcomes but rather that we should look at how experimental work *introduces* new constraints: "As features of the instruments, theories, and procedures are better understood, the number of constraints on interpretation increases" (Galison 1987: 132).

Hacking follows the same path to a certain extent, stressing – especially in the works after *Representing and Intervening* – how experimentation introduces new constraints on how phenomena can be interpreted. Drawing insights from multiple sources, including Galison's idea that not only experiments but also "instruments have a life of their own" and Pickering's extension of Duhem's

thesis about auxiliary hypotheses, Hacking highlights the interplay between several levels of "plastic resources", including "theory, phenomenology of the apparatus, and the material instrumentation and objects being investigated" (1991: 237). The mutual adjustments among these levels results in both the stability (Hacking 1988b) and the "self-vindicating feature" of laboratory sciences (Hacking 1992b). When we look at Hacking's work in further detail, more specific debates come to the foreground – the realism-antirealism debate (4.2.1), the Science Wars (4.2.2), the debate on incommensurability (4.2.3) – but these nonetheless fit into the same program.

4.1.1 Experimentation as a new form of realism

Another important piece of context for Hacking's work is the realism-antirealism debate of the 1970s and 1980s, which centered around questions concerning the nature and function of scientific theories and theoretical entities: are theories objective descriptions of an independent reality or mere instrumental tools to make predictions? Do theoretical entities such as atoms really exist or are they just useful fictions? Whereas scientific realists believed that entities, states, and processes described by true theories referred to genuine entities in the world (Putnam 1971), scientific anti-realists denied this (van Fraassen 1980). For the anti-realists, scientific theories were instead instruments that at best could be useful or apt but not 'true' in the ordinary sense (van Fraassen 1980: 88).

Developed against the backdrop of this debate, Hacking's *Representing and Intervening* can be read as a reply to van Fraassen's constructive empiricism and to the corresponding claim that "science aims to give us theories which are empirically adequate; and acceptance of a theory involves a belief only that it is empirically adequate" (van Fraassen 1980: 12). Here Hacking introduces an idea that is typically linked to discussions of his philosophy of experiment: *experimental realism*. Whereas van Fraassen's anti-realist stance encompasses both theories and theoretical entities, Hacking argues, on the contrary, that in many cases we can have compelling evidence supporting our belief in the existence of electrons without necessarily having a plausible theoretical description of them. This evidence is provided by our ability to manipulate theoretical, postulated entities and use them to intervene in causal nexuses in the world. "We shall count as real what we can use to intervene in the world to affect something else, or what the world can use to affect us" (Hacking 1983: 146).

Hacking is thus anti-realist about theories but realist about theoretical entities. He grounds this distinction in the claim that "[i]f you can spray electrons, then they are real" (Hacking 1983: 23). In other words, more than

the theory describing the electron, it is the scientist's ability to "shoot them" with a polarizing electron gun (in an experiment measuring the charge of the quark) that provides the best evidence of their actual existence. His point is that "by the time we can use the electron to manipulate other parts of nature in a systematic way, the electron has ceased to be something hypothetical, something inferred" (Hacking 1983: 262). He thus mobilizes philosophy of experiment to shift the existing realism-antirealism debate, by introducing a new kind of realism whose defining traits are the active, pragmatic, and heuristic functions of experimentation.

4.1.2 The science wars

The realism vs. antirealism debate eventually gave birth to a subsequent debate in the 1990s known as the Science Wars, which opposed scientists and constructivists (Ross 1996; Gieryn 1999). One of the main points of divergence or "sticking points" dividing the two sides was whether scientific classifications are "natural", that is, found in nature or humanly created (Hacking 1999).

In this context, Hacking was able to draw on the distinction he had previously developed between natural kinds and interactive kinds (Hacking 1995). According to Hacking, the objects of the natural sciences are natural kinds and indifferent to our categorization, whereas human or social kinds are shaped by the interaction between a given scientific category and a subject thus categorized. This divide partially maps onto the Hacking distinction we already encountered in 3.2, between facts that are historically *contingent* and those that are historically *constituted*. Hacking uses the examples of the laser and the maser (Hacking 1983: 226-227; 1999: 179-180), phenomena that might not have occurred in the universe before we created them. In that sense Hacking identifies them as historically contingent. Yet this does not mean they are unreal or not objective. In fact, far from being historically or humanly constituted, lasers and masers for Hacking are natural kinds.

On the other hand, within a constructivist framework, the historical and technical context of the production of a given phenomenon represents its own condition of existence. A constructivist like Latour would argue that all theoretical entities are historically contingent and constituted, both those belonging to the natural sciences and those of the social sciences. For Hacking, here constructivists are overplaying their cards so to speak. They go too far in their claims about natural kinds, though he concedes that they do have a point concerning the constructed aspect of classifications such as autism or homosexuality. The claim about entities being historically constituted is thus not completely wrong, but Hacking argues that this is only correct for the social sciences, whose theoretical objects, being both historically contingent and

historically constituted, did not exist in any specifiable form until they become objects of scientific inquiry (Hacking 2002: 11).

Thus once again, Hacking mobilizes elements of his philosophy of experiment to dismantle existing debates and move them in new directions.

4.1.3 Experimentation as new form of continuism

A final debate that Hacking tried to shake up was the one over the incommensurability of science initiated by Thomas Kuhn and Paul Feyerabend in the 1970s (Kuhn 1962; 1967; Feyerabend 1975). This is true not only of Hacking but also of Galison. Indeed, it is possible to see the emergence of the philosophy of experiment in the 1980s as a direct response to the thesis of scientific incommensurability.

In 3.2 we mentioned that, for Hacking, stable laboratory phenomena are resistant to changes in theory and can cross through and accumulate through different theoretical frameworks. Seen in this light, experimental techniques and results are the best ways to assess scientific progress. Hacking locates new forms of continuity in the production of phenomena and experimental styles. Phenomena and instruments have a "life of their own", as they have the tendency to accumulate through theoretical changes. Hacking links this insight to his interest in the different styles of scientific reasoning, ranging from mathematical postulation to statistics, which according to Hacking are also accumulative:

> What we accumulate are *experimental techniques* and *styles of reasoning.* Anglophone philosophy of science has too much debated the question of whether theoretical *knowledge* accumulates. Maybe it does not. So what? Phenomena and reasons accumulate. (Hacking 2002: 45)

This passage presents the same theme that is also present in Galison's critical postmodern model described in 2.2: a complex and patchy vision suggesting that there is "no single way in which the patterns of continuity and discontinuity are aligned and there is no reductive hierarchy" (Galison 1988: 209). The Galison model aims to recuperate the best of both the positivist and antipositivist models, while incorporating new insights drawn from new experimentalism:

> By breaking up the experimental level into intercalated patches of continuity and discontinuity we incorporate the insight of the antipositivists: experiment and experience do not give unmediated access to universal, basic propositions. At the same time, by allowing experiment to continue across theoretical breaks, we (partially) resurrects the positivists' contention that theories do change while leaving unbroken a chain – or at least a surviving. (Galison 1988: 209)

Thus according to Galison, there can be independent continuity in experimental knowledge despite clear theoretical breaks: "periods during which theorists break with tradition do not necessarily correspond to disruptions in the subject matter, methods, procedures, and instruments of experimental physics" (Galison 1987: 13). Countering the incommensurability thesis, Galison points out that "experimental conclusions have a stubbornness not easily cancelled by theory change" (Galison 1987: 259).

Yet, Hacking's and Galison's positions do not entirely overlap. According to Hacking, for instance, it is not only phenomena that accumulate through theoretical shifts but also statements of observation. Hacking repeats the neopositivists' belief that observational statements are made in a pre-theoretical language which makes translation between the two different theories possible (Hacking 1983: 167-185). Furthermore, he sees both neo- and anti- or post-positivists philosophers and historians of science as sitting on the same side of the divide between theory and experimentation: their approach remain theoretical and incapable of accounting for the relative autonomy of experiments. This ancillary role assigned to experimentation with respect to theory in the so-called "standard image" of science was not subverted by anti-positivists such as Karl Popper who, as Hacking remarks, believed the experimenter should not begin work until the theoretician has finished their job (Hacking 2008: 109).

Despite these differences, Hacking and Galison seem to agree on the fact that elements of a philosophy of experiment has the potential to steer the old debate over incommensurability into new directions.

5. *Conclusion*

Our reconstruction has shown that experimentation did not appear out of the blue in philosophy of science in the 1980s. A variety of traditions reflecting on the role of experiments in science existed before and were still operational at that time. We have also demonstrated that when these earlier traditions are mentioned by more recent scholars, they are often not simply 'rediscovered'. The contextualist narrative we put forward (§4) highlights how specific versions of both the invention (§2) and the rediscovery (§3) narratives that philosophers of experiment articulate usually reveal their goals vis-à-vis larger debates: there are always specific reasons why philosophers of experiment claim to either have invented the philosophical topic of experimentation or rediscovered older philosophies of experiment.

As a consequence, we maintain that experimentation should not be considered a fixed notion that is either present or not in various philosophical, historical, and sociological discourses. Rather, in its myriad appearances the

topic can be grouped in (at least) two important ways: (a) by what is meant by experimentation and (b) by what the philosopher in question wishes to achieve by invoking experimentation. Only a contextualist narrative can resolve the problems we encountered with both the invention and the rediscovery narratives while simultaneously acknowledging the specificity of Hacking's answers to both (a) and (b).

Of course, if one wishes to assess Hacking's contribution to the history of the philosophy of experiment fully, a great number of other questions remain. Once one accepts that experimentation can shift in meaning and use, new avenues for a genuine history of the philosophy of experiment are opened up. We therefore would like to end this article by briefly highlighting three different avenues for future research in the history of the philosophy of experiment.

First of all, we must take into account shifts that have occurred in *philosophy* itself, not only in terms of its professionalization but also in terms of the specific topics and issues it deems central or relevant. From this perspective, when assessing recent or earlier philosophies of experiment, it is key to understand how experiments emerged for these authors as a site of philosophical preoccupation. For instance, to assess Hacking's innovative contribution to the philosophy of experiment, we first must identify and evaluate what motivated his reflections on experimentation. This is the approach we have tried to employ in this paper, which we think could be extended to other recent philosophers of experiment.

Secondly, we must examine the '*regional' meaning of experimentation in different domains of knowledge* as well as *developments within these sciences* themselves – not only in terms of how specific scientific disciplines have developed throughout the 20th century but also in terms of which science is considered paradigmatic at specific moments in time. From this perspective, it is to be expected that reflecting on the form of mathematical physics dominant in the period immediately following Einstein's theories, as in Bachelard's case, would require a very different approach than working on the high-energy physics and biochemistry prevalent in the second half of the 20th century that Hacking and Latour respectively studied.

Finally, to fully grasp Hacking's philosophy of experiment also requires examining how the *role of experimentation itself has shifted in science and society*. From this perspective, one might wonder, for instance, whether the new philosophical interest in experimentation in the 1980s was in part a product of shifts in the institutional structure of science – part of developments that put instruments and intervention in the spotlight, such as Big Science (Pestre 1997) and the commodification of science (Lyotard 1979) – in ways that would not have applied for Bachelard or Dingler.

All of these questions deserve detailed analysis. While in this paper we have limited ourselves to the first, philosophical question, our hope is that future research will help elaborate a history of the philosophy of experiment that addresses all three – and gives authors such Hacking the places in that history that they rightfully deserve.

Matteo Vagelli
matteo.vagelli@univ-paris1.fr
Centre de philosophie contemporaine de la Sorbonne,
Université Paris 1 Panthéon-Sorbonne

Massimliano Simons
massimiliano.simons@ugent.be
Department of Philosophy and Moral Sciences, Ghent University

References

Ackermann, Robert, 1985, *Data, Instruments, and Theory: A Dialectical Approach to Understanding Science*, Princeton University Press, Princeton.

Ackermann, Robert, 1989, "The New Experimentalism (Book Review)", in *The British Journal for the Philosophy of Science*, 40, 2: 185-190.

Agazzi, Evandro, Gerthard Heinzmann, 2015, eds., *The Practical Turn in Philosophy of Science*, Franco Angeli, Milan.

Arabatzis, Theodore, 2008, "Experiment", in Stathis Psillo and Martin Curd, eds., *The Routledge Companion to the Philosophy of Science*, Routledge, London: 159-170.

Bachelard, Gaston, 1927, *Essai sur la connaissance approchée*, Vrin, Paris.

Bachelard, Gaston, 1938, *La formation de l'esprit scientifique*, PUF, Paris; Eng. tr. by Mary MacAllester Jones 2002, *The Formation of the Scientific Mind*, Clinamen Press, Manchester.

Bachelard, Gaston, 1949, *Le rationalisme appliqué*, PUF, Paris.

Bachelard, Gaston, 1951, *L'activité rationaliste de la physique contemporaine*, PUF, Paris.

Bachelard, Gaston, 1953, *Le matérialisme rationnel*, PUF, Paris.

Bachelard, Gaston, 1931 [2005], "Noumena and Microphysics", in *Angelaki: Journal of Theoretical Humanities*, 10, 2: 73-78.

Bontems, Vincent, 2010, *Bachelard,* Belles Lettres, Paris.

Burian, Richard, 2007, "On microRNA and the Need for Exploratory Experimentation in Post-Genomic Molecular Biology", in *History and Philosophy of the Life Sciences*, 29, 3: 285-311.

Canguilhem, Georges, 1955, *La formation du concept de réflexe au XVII et XVIII siècles*, Vrin, Paris.

Dagognet, François, 1965, G*aston Bachelard: Sa vie, son œuvre, avec un exposé de sa philosophie*, PUF, Paris.

Dagognet, François, 1979, *Mémoire pour l'avenir : Vers une méthodologie de l'informatique*, Vrin, Paris.

Dingler, Hugo, 1928, *Das Experiment: Sein Wesen und seine Geschichte*, Reinhardt, München.

Feest, Uljana, Friedrich Steinle, 2014, „Experiment", in Paul Humphreys, ed., *The Oxford Handbook of Philosophy of Science*, Oxford University Press, New York: 274-295.

Feyerabend, Paul, 1975, *Against Method*, Verso, London.

Franklin, Allan, 1986, *The Neglect of Experiment*, Cambridge University Press, Cambridge.

Galison, Peter, 1987, *How Experiments End*, University of Chicago Press, Chicago.

Galison, Peter, 1988, "History, Philosophy, and the Central Metaphor", in *Science in Context*, 2, 1: 197-212.

Galison, Peter, 1997, *Image and Logic: A Material Culture of Microphysics*, University of Chicago Press, Chicago.

Gooding, David, Trevor Pinch, 1989, *The Uses of Experiment: Studies in the Natural Sciences*, Cambridge University Press, Cambridge.

Gieryn, Thomas, 1999, *Cultural Boundaries of Science: Credibility on the Line*, University of Chicago Press, Chicago.

Hacking, Ian, 1982, "Experimentation and Scientific Realism", in *Philosophical Topics*, 13, 1: 71-87.

Hacking, Ian, 1983, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge University Press, Cambridge.

Hacking, Ian, 1988a, "Philosophers of Experiment", in *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2: 147-156.

Hacking, Ian, 1988b, "On the Stability of the Laboratory Sciences", in *The Journal of Philosophy*, 85, 10: 507-514.

Hacking, Ian, 1992a, "Ian Hacking: An Interview by R. B. De Sousa", in *Toronto Philosophy News*, 1, 2: 4-6.

Hacking, Ian, 1992b, "The Self-Vindication of the Laboratory Sciences", in Andrew Pickering, ed., *Science as Practice and Culture*, University of Chicago Press, Chicago: 29-64.

Hacking, Ian, 1995, "Looping Effects on Human Kinds", in Dan Sperber, David Premack and Ann James Premack, eds., *Causal Cognition: A Multidisciplinary Debate*, Clarendon Press, Oxford: 351-394.

Hacking, Ian, 1999, *The Social Construction of What?*, Harvard University Press, Cambridge.

Hacking, Ian, 2002, *Historical ontology*, Harvard University Press, Cambridge.

Hacking, Ian, 2006, «La philosophie de l'expérience : illustrations de l'ultrafroid», *Tracés*, 11, 1: 195-228.

Hacking, Ian, 2008, *Scientific Reason*, Taiwan University Press, Taipei.

Knorr-Cetina, Karin, 1981, *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Pergamon, Oxford.

Kuhn, Thomas, 1962, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Kuhn, Thomas, 1977, *The Essential Tension*, University of Chicago Press, Chicago.

Latour, Bruno, Steve Woolgar, 1979, *Laboratory Life: The Social Construction of Scientific Facts*, Sage, Beverly Hills.

Lenhard, Johannes, 2007, "Computer Simulation: The Cooperation between Experimenting and Modeling", in *Philosophy of Science*, 74, 2: 176-194.

Lynch, Michael, 1985, *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*, Routledge and Kegan Paul, Boston.

Lyotard, Jean-François, 1979, *La condition postmoderne: Rapport sur le savoir*, Editions de Minuit, Paris.

Mayo, Deborah, 1996, *Error and the Growth of Experimental Knowledge*, University of Chicago Press, Chicago.

Pestre, Dominique, 1997, "Science, Political Power and the State", in John Krige, Dominique Pestre, eds., *Science in the Twentieth Century*, Harwood, Reading: 61-76.

Pickering, Andrew, 1984, *Constructing quarks: A sociological history of particle physics*, Edinburgh University Press, Edinburgh.

Putnam, Hilary, 1971, *Philosophy of Logic*, Allen & Unwin, London.

Radder, Hans, 1993, "Science, Realization and Reality: The Fundamental Issues", in *Studies in History and Philosophy of Science*, 24, 3: 327-349.

Radder, Hans, 2009, "The Philosophy of Scientific Experimentation: a Review", in *Automated Experimentation*, 1, 2: 1-8.

Radder, Hans, 2003, ed., *The Philosophy of Scientific Experimentatio*n, University of Pittsburgh Press, Pittsburgh.

Ravetz, Jerome, 1971, *Scientific Knowledge and its Social Problems*, Claredon Press, Oxford.

Ross, Andrew, 1996, ed., *Science Wars*, Duke University Press, Durham.

Schatzki, Theodore, Karin Knorr-Cetina, Eike von Savigny, 2001, eds., *The Practice Turn in Contemporary Theory*, Routledge, London/New York.

Shapin, Steven, Simon Schaffer, 1985, *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*, Princeton University Press, Princeton.

Simons, Massimiliano, 2018, "The Janus head of Bachelard's phenomenotechnique: From purification to proliferation and back", in *European Journal for Philosophy of Science*, 8, 3: 689-707.

Soler, Léna, Sjoerd Zwart, Michael Lynch, Vincent Irsrael-Jost, 2014, eds., *Science after the Practice Turn in the Philosophy, History, and Social Studies of Science*, Routledge, London/New York.

Steinle, Friedrich, 2002, "Experiments in History and Philosophy of Science", in *Perspectives on Science,* 10, 4: 408-432.

Traweek, Sharon, 1988, *Beamtimes and Lifetimes: The World of High Energy Physicists,* Harvard University Press, Cambridge.

Vagelli, Matteo, 2014, "Ian Hacking, The Philosopher of the Present. An Interview by Matteo Vagelli", *in Iride. Journal of Philosophy and Public Debate*, 27, 72: 239-269.

Vagelli, Matteo, 2017, "Bachelard, Hacking e il realismo tecnoscientifico", in Paolo Donatiello, Francesco Galofaro, Gerardo Ienna, eds., *Il senso della tecnica. Saggi su Bachelard*, Esculapio, Bologna: 121-136.

Van Fraassen, Bas, 1980, *The Scientific Image*, Oxford University Press, Oxford.

Winsberg, Eric, 2009, "Computer Simulation and the Philosophy of Science", in *Philosophy Compass*, 4, 5: 835-845.

Understanding stability in cognitive neuroscience through Hacking's lens[1]

Jacqueline Sullivan

Abstract: Ian Hacking instigated a revolution in 20th century philosophy of science by putting experiments ("interventions") at the top of a philosophical agenda that historically had focused nearly exclusively on representations ("theories"). In this paper, I focus on a set of conceptual tools Hacking (1992) put forward to understand how laboratory sciences become stable and to explain what such stability meant for the prospects of unity of science and kind discovery in experimental science. I first use Hacking's tools to understand sources of instability and disunity in rodent behavioral neuroscience. I then use them to understand recent grass-roots collaborative initiatives aimed at establishing stability in this research area and tease out some implications for unity of science and kind creation and discovery in cognitive neuroscience.

Keywords: cognition; laboratory science; natural kinds; stability; unity of science

A primary aim of cognitive neuroscience is to understand the neural underpinnings of human cognition. Cognitive neuroscience roughly may be divided into two branches, one which focuses on humans and human clinical populations, and the other, which focuses on non-human animals (e.g., non-human primates, rodents). In this paper, I will be concerned primarily with rodent behavioral neuroscience. Intervention experiments in rodents are crucial for identifying the neural mechanisms that give rise to cognition in humans; rodents afford the possibility of using state-of-the-art techniques to alter genetic, molecular, or circuit-level activity and determine the impact of such manipulations on behavioral performance in tasks designed to assess human relevant cognitive functions. Consider a cognitive function like decision-making, in which an organism has to make a choice between two available actions. A rodent behavioral researcher may design a task to assess decision-making in mice and artificially alter the activity of a population of neurons (e.g., dopamine neu-

[1] The author would like to thank Matteo Vagelli and Marica Setaro for helpful comments on an earlier draft of this paper.

ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.346

rons in the medial striatum) *in vivo* as mice perform the task in order to assess the impact of this manipulation on the mouse's performance. The same basic approach may be used to investigate a range of cognitive functions including: *working memory*, *cognitive flexibility*, *attention*, *motivation* and *response inhibition*, to name only a handful. Insofar as rodent behavioral research occurs in laboratories and involves the use of "apparatus used in isolation to interfere" (1992: 34) and, as I will show, to "create new phenomena", it may be regarded as constituting a laboratory science in Hacking's sense.[2]

During the past two decades, a picture has emerged in philosophy of science with respect to how areas of neuroscience directed at understanding the neural underpinnings of cognition, like rodent behavioral neuroscience, make progress (e.g., Bechtel 2008; Craver 2007; Picinnini & Craver 2011). The basic idea is that neuroscientists seek multi-level mechanistic explanations that describe the physical entities/components (e.g., neurons, neural circuits) and activities/processes (e.g., neuronal firing, neurotransmitter release) that bring phenomena of interest (e.g., reward-based learning) about. To take a widely cited example from the philosophical literature, activation of N-methyl-D aspartate receptors in area CA1 of the rat hippocampus is one component in the description of the multi-level mechanism of spatial memory (e.g., Craver 2007). Progress in neuroscience on the mechanistic view occurs as findings from experiments being undertaken in the same and different areas of neuroscience are "seamlessly integrated" into descriptions of multi-level mechanisms of cognitive phenomena (e.g., Picinnini & Craver 2011).

This view of progress in the mind-brain sciences has arisen primarily from the perspective of the philosophy of explanation, in the absence of careful evaluation of the precise kind of knowledge that individual neuroscientific experiments and research studies yield and absent an analysis of how results from different studies actually fit together. In the last two decades of the 20th century, however, Hacking (e.g. 1983; 1991; 1992) urged philosophers of science to relinquish their exclusive focus on "representations" (i.e., theories, explanations) and turn their attention to experiments ("interventions") – those processes by which phenomena are produced or "created" in scientific labo-

[2] Hacking (1992: 37) is not interested in "research at the frontiers of inquiry", which "can be as unstable as you please". Rodent behavioral neuroscience, insofar as it employs cutting-edge tools, is thus not an example of a stable science. My aim in this paper, however, is to use Hacking's tools to identify features of rodent behavioral neuroscience that investigators themselves believe jeopardize the production of stable knowledge – knowledge that they regard as necessary for progress in their field. I aim to show that Hacking's tools may be used to characterize the kind of stable knowledge this area of science lacks and that some scientists working in this area seek. As I see it, Hacking's descriptive conceptual tools may be prescriptively used to understand how to stabilize laboratory sciences (even if he did not intend them to be used in this way).

ratories. In this paper, I use conceptual tools that Hacking put forward to evaluate experimental practices in rodent behavioral neuroscience. I focus primarily on a set of conceptual tools that Hacking (1991; 1992) put forward to understand the stability of the laboratory sciences, which he used to tease out implications for the prospects of unity of science and kind discovery in experimental science.

I begin, in Section 1, by identifying some preliminary features of rodent behavioral neuroscience. The aim is to provide enough detail that Hacking's taxonomic framework for understanding stability in experimental science may later be applied. I then consider Hacking's (1991) claim that disunity is not a temporary feature of science, but, indeed, a permanent aspect of the scientific landscape, because, despite Thomas Kuhn's (1961) claims about revolutions in science, science does accumulate, and some parts do become stable – –a kind of local stability that is antithetical to global unity of science. I then describe the conceptual tools that Hacking (1992) introduced and used in order to understand how laboratory science becomes stable, and I use these tools for two purposes. First, I apply them to characterize the sources of current instability in rodent behavioral neuroscience. Second, I use them to illuminate strategies of stabilization currently being collaboratively implemented in the context of two grass-roots initiatives in this field. I conclude by teasing out implications pertaining to unity of neuroscience and the nature of the kinds that neuroscience, on Hacking's view, is likely to discover.

1. *Some preliminary observations about rodent behavioral research*

In neuroscience, rodent behavioral studies may be aimed at understanding species-specific cognition (e.g., mouse cognition), but rodents are more commonly used as models for humans – mammalian "stand-ins" that afford the possibility of combining tasks to assess cognition with cutting-edge visualization (e.g., fiber photometry) and intervention (e.g., optogenetics) technologies for intervening in molecular, cellular, and neural circuit activity to determine the impact of such interventions on cognitive functioning. Mouse models of neuropsychiatric, neurodegenerative and other brain disorders (e.g., concussion) also are regarded as crucial for identifying the neural mechanisms that underlie impairments in cognitive functions such as memory, attention and decision-making that accompany these disorders and developing effective therapeutic interventions to treat them.

One cognitive function that is crucial for navigating the world on a daily basis, and that is impaired in neurodegenerative diseases like Alzheimer's disease, is *pattern separation* – namely, the ability to distinguish memories from

each other, to separate one memory pattern from the next. Consider a simple illustrative example (Beckinschtein, Kent, Ooman, *et al.*, 2013). If you drive a car to school or the office, it is likely that you park your car in a different spot each day. Yet, you typically are able to remember where you parked your car in the parking lot each day, despite parking in a different spot from day to day. This is an instance of pattern separation.

One task that neuroscientists have used to study pattern separation in rodents is the Spontaneous Location Recognition (SLR) task. In one version of this task, a rodent is placed in an open circular arena and allowed to habituate to that environment. Three novel stimuli (different objects) are then introduced. In a simple version of the task, there is a sample phase in which two of the objects are placed 50° apart from each other and the third object is placed equidistant from each of the other two. The mouse is then placed into the arena to explore. In the choice phase of the task, which occurs 24 hours after the sample phase, two novel copies of the two objects that were placed 50° apart during the sample phase are presented. This time, however, one of the two objects is placed in a novel location (a location equidistant between its previous location and the "familiar" location of the other object).[3] Given that mice prefer novelty, a mouse that spends more time exploring the object in the novel location compared to the object in the familiar location divided by the total time it spends exploring is considered to have successfully "pattern separated" – i.e., to have a memory that enables them to distinguish the object in the novel location from the object in the familiar one (See Beckinschtein, Kent, Ooman, *et al.* 2013).

Rodent behavioral neuroscience is an interdisciplinary area of science that brings together investigators hailing from a variety of different fields including: genetics, animal behavior, neurophysiology, biochemistry and computational neuroscience, to name only a handful. Tasks like SLR may be combined with a variety of different visualization and intervention techniques that allow the activity of molecules, cells, and circuits to be detected and manipulated to determine the impact on behavioral performance. Mouse models of neuropsychiatric and neurodegenerative disease and other brain disorders (e.g., concussion) may be used in order to identify disruptions in neural circuit activity that underlie impaired performance on such tasks. There are many different apparatuses (e.g., mazes, open fields, classical conditioning chambers, touchscreen operant chambers (described in section 4)) and tasks (e.g., SLR, contextual fear conditioning, social recognition, paired associates learning) that rodent

[3] As rodents have a keen sense of smell, new, identical objects are used in the choice phase to rule out the possibility that the rodents are using olfactory cues to perform the task.

behavioral researchers may use in combination with intervention techniques to investigate the neural bases of different kinds of cognitive functions.

With this brief introduction to rodent behavioral neuroscience, I turn now to Hacking's views about disunity of science.

2. *Disunified sciences*

In "Disunified Sciences" (1991), Hacking identifies and evaluates a set of eleven theses characteristic of logical positivist understandings of the unity of science. I want to briefly consider a relevant subset of these, given that certain aspects of unity of science conceived of by the logical positivists remain implicit in contemporary thinking about progress in areas of neuroscience like rodent behavioral neuroscience. Specifically, recent arguments for unification focus on explanatory integration, which involves the integration of data from multiple experiments into explanations of cognitive functions (e.g., Craver 2007; Craver, Piccinini 2011). Yet, such integration is only possible if the constructs designating cognitive functions under which data from different experiments are being integrated are stable. As I will argue later in the paper, there are good reasons to think they are not.

Among the unity theses that Hacking (1991: 41) considers are two "metaphysical theses": (1) there is a single world, and it contains diverse kinds of phenomena that are (2) "interconnected". The epistemic aim of science is to understand this single world and science offers the best method for attaining such understanding. The logical positivists expressed confidence in the idea that "there is one right fundamental system of classifying everything" (taxonomic thesis), that will be expressed in a single scientific language (e.g., physics) (linguistic thesis) that identifies the stable regularities and tracks so-called "natural kinds" (Hacking 1991: 41). They believed that science gradually approximates towards this one right system by means of intertheoretic reduction (reductionist thesis)-namely, the establishment of bridge laws, as terms in reducing and reduced theories are connected (connectability), and reducing theories come to explain all the phenomena originally explained by the reduced theory (derivability). In the process, unity of science is achieved, as "many facts" are brought "under the wing of one intellectual structure" (Hacking 1991: 41). Moreover, the achievement of unity was not only descriptively accurate with respect to the history of science, but also an "on-going trend" in the heyday of logical positivism (Oppenheim, Putnam 1958).

Hacking aims to demonstrate that none of these unity theses are applicable to contemporary science. I am particularly interested in his arguments against intertheoretic reduction and the discovery of a single system of scientific clas-

sification that tracks natural kinds. Hacking offers two different lines of argument here. First, he notes the sheer diversity in language and methods that we find in contemporary science and the difficulties that heads of academic departments face in trying to unify different areas of science that may generally be classified as, for example, "biological", within a single "super-department" (Hacking 1991: 43). He points out that even Oppenheim and Putnam, who advocated for unity via theory reduction, and described it as an ongoing trend in science, themselves acknowledged certain "incompatible trends" in science that were antithetical to unity. Hacking also points to how "overspecialized" science has become to the extent that "in a quite straightforward sense there is no common language of science, and [. . .] as a matter of practicability, there could not be" (Hacking, 1991: 44). Yet, Hacking, agreeing with philosopher of science Patrick Suppes, does not regard "the irreducible pluralism of languages of science" as an obstacle to "the continued growth of science" (Suppes 1984: 121 as quoted in Hacking 1991: 44).

Hacking's second related strategy for establishing the disunity of science is to argue that much of contemporary science, rather than moving towards theoretical unity, becomes stable within a restricted domain. He relies in part on examples from physics to support this claim. For example, he points to scientists like Sheldon Glashow and Werner Heisenberg who have described Newtonian mechanics and classical quantum mechanics as theories that are not universally true, but "valid in [their] domain" (Hacking 1991: 48). Hacking notes that "the idea of a closed theory with its domain at once suggests disunity: different domains governed by different theories" (Hacking 1991: 48) rather than theory displacement or theory reduction. Hacking insists that from the vantage point of philosophy of scientific experimentation, we encounter a similar kind of stability when we look at laboratory science; "[it] is stable" he claims, "not because there is a domain of experiment, given by nature itself, to which certain theories are true" but "because there is a mutual maturing of types of apparatus, phenomena and theory" (Hacking 1991: 49). Such stability results in disunity, in part, because each laboratory science constitutes its own domain in which "bodies of knowledge" are not discarded but rather "supplement[ed] with new kinds of instruments" (Hacking 1991: 49).

At the end of "Disunified Sciences", Hacking emphasizes the need for a more detailed set of conceptual tools to analyze laboratory sciences and to understand how "experimental stability" emerges. He provides one such set of conceptual tools in "The Self-Vindication of the Laboratory Sciences" (1992), which is the focus of the next section.

3. *Hacking's lens: the view from the philosophy of experiment*

Hacking (1992) acknowledges that he regards his thesis about the stability of laboratory sciences as "an extension of [Pierre] Duhem's [coherentist] doctrine, that a theory [shown to be] inconsistent with an observation can always be saved by modifying an auxiliary hypothesis", even a hypothesis about the working of the very instrument used to test the theory (1992: 30). However, according to Hacking, Duhem's framework for understanding what happens in experimental contexts is inadequate because it focuses only on how stability is achieved as representations – "theory and auxiliary hypothesis" are "adjusted to each other". Hacking believes philosophers require a richer and more diverse set of tools for understanding experimentation and the stability of experimental science that can accommodate how much of the knowledge generated by the laboratory sciences is stable and how the devices and practices become permanent fixtures of the scientific landscape. To this end, he puts forward a "taxonomy of elements of experiment which [he claims] are mutually adjusted" or brought into coherence so as "to produce the self-vindicating character of laboratory science" (Hacking 1992: 32). He divides these elements into three categories that are intentionally broad so that each captures a wide range of items: (1) *ideas*, (2) *things* and (3) *marks*. In the rest of this section, I will consider each of these in turn.

Although Hacking sought to shift emphasis in philosophy of science away from theories and towards experiments, he did recognize the role that "representations" play in experimental contexts. The category of "Ideas" includes those empirical questions about a phenomenon of interest that an experiment is designed to answer. For example, is activation of a specific population of neurons necessary for spatial memory or visual associative learning? Questions may also be directed at the merits and failings of large-scale scientific theories, which is particularly common in areas of science like physics, but uncommon in areas like rodent behavioral neuroscience. Background knowledge or background beliefs on which an investigator relies, which may be neither systematized nor made explicit also fall into Hacking's *ideas* category. Background beliefs could range from an investigator's understanding of a concept such as *spatial memory*, to her understanding of how a given intervention or visualization technique (e.g., optogenetics) works, to her assumptions about potential confounds to be controlled for during an experiment (e.g., feeding times for a rat subject when successful task performance requires hunger as motivation). "Ideas" also include high-level "systematic theories" about the subject matter under study and "topical hypotheses" that are local to experimental contexts and connect together theoretical ideas with the implementation of those ideas

in the laboratory in a way that is revisable. A final element within the category of ideas involves the understanding on the part of the investigator as to the nature and structure of the apparatus (e.g., task analysis) or tools that are used to produce data and how those tools actually work.

Hacking's second category, "things" – includes all of the material elements involved in an experiment such as: the targets of investigation (e.g., mice and rats, cells, molecules, synapses) and the instruments or apparatuses (e.g., optogenetic techniques) used "to alter or interfere" with those targets (Hacking 1992: 46). The instruments that serve a productive function, for Hacking, insofar as they are used to "create phenomena" (e.g., Hacking 1983; 1992) differ from those instruments that are used to detect the effects of the intervention – to "determine or measure the result of the interference or modification of the target" (Hacking 1992: 47). The broader category of "*tools*" consists of "all the humble things upon which the experimenter must rely" in order to run the experiment – for example, microtomes for slicing tissue samples, artificial cerebrospinal fluid for preserving brain tissue samples, or the computer equipment and software for running a given cognitive task. Finally, *Data generators* are the parts of the experiment that generate the data (Hacking 1992: 48), such as movement tracking devices and reaction-time software–all of the programs that record data, including scientists recording data by hand.

Hacking's final category, "marks and the manipulation of marks" is intended to include the outputs of experiments – the data – as well as those processes to which the data are subject. In order to be interpretable, data must be reduced and analyzed statistically. Yet, Hacking remarks that it is important to remember that choice of data reduction, data enhancement and data analysis techniques are often influenced by "ideas" on the part of investigators including background knowledge and theoretical commitments. The final interpretation of the data is also done in light of the researcher or research team's background knowledge, understanding of how the apparatus and other tools used in the experiment work and, where relevant, high-level theory. Hacking claims that an important part at this stage of the laboratory work is an "estimation of systematic error, which requires explicit knowledge of the theory of the apparatus –and which has been too little studied by philosophers of science" (Hacking 1992: 49). Since the publication of Hacking's paper, a number of philosophers of science have sought to fill this gap (e.g. Mayo 1991; 1996; Sullivan 2018; Schickore 2005; 2019).

According to Hacking, the stability of a laboratory science is gradually established as these 15 elements falling into the broader categories of "thoughts, actions, materials, and marks" are "mutually adjusted to each other" and "what meshes (Kuhn's word) is at most a network of theories, models, approxima-

tions, together with understandings of the workings of our instruments and apparatus" (Hacking 1992: 30). Laboratory sciences become self-vindicating on Hacking's picture, insofar as eventually, "any test of theory is against apparatus that has evolved in conjunction with it – and in conjunction with modes of data analysis" (1992: 30).

Importantly, laboratory scientists have to engage in strategies of stabilization that bring these different elements into consilience. Although Hacking does not acknowledge it explicitly, laboratory sciences do not consist of a single laboratory running experiments in isolation, but investigators – research teams – running experiments in many different laboratories. The stability of experimental science that Hacking describes is thus not something that comes about in a single laboratory, but rather, across many different laboratories having investigators who share *thoughts*, *actions, materials*, *marks and strategies for manipulating marks* in common and who are collaboratively united in bringing these elements into productive symbiosis.

As I aim to show in the next two sections, Hacking's taxonomy of elements and views about the stability of laboratory science may be used as a foil for understanding why instability may occur in some laboratory sciences, not merely due to the fact that these sciences are on the cutting-edge, but also that researchers in the field may be engaged in practices that effectively destabilize the field insofar as their actions are not directed at bringing these elements into consilience. In such instances of instability, we may anticipate a lack of conceptual, methodological and explanatory unity within these fields. Also, in light of Hacking's framework, the possibility that a given laboratory science may stabilize in any number of ways depending upon who the actors are, and what ideas, actions, materials and marks they aim to bring into consilience is consistent with local unity, but as Hacking (1991) indicates, global disunity.

In the next section (Section 4), I use Hacking's framework to identify those aspects of experimental practice in rodent behavioral neuroscience that have served to promote the instability of the field and have been a barrier to the production of stable knowledge pertaining to the neural underpinnings of rodent cognition. The kind of instability that we encounter here is consistent with what might be regarded as counterproductive disunity. Yet, if Hacking is correct, there is such a thing as productive disunity – and it correlates with areas of science implementing strategies to arrive at stable knowledge – strategies that simultaneously result in the creation of phenomena, and the development of specialized languages and methods and associated practices that co-evolve and become "self-vindicating".

4. *Rodent behavioral neuroscience through Hacking's lens*

In Section 1, I briefly described some basic features of the structure of experiments in rodent behavioral neuroscience. I now want to elaborate on the structure of research in this field and evaluate it by way of Hacking's framework of "ideas, things, and marks".

First, consider Hacking's concept of "ideas", a category which includes empirical questions about phenomena of interest, high-level theories, background assumptions, topical hypothesis that relate theories to observations made in experimental contexts and beliefs about how a given experimental apparatus or tool works. As I mentioned in Section 1, researchers working in rodent behavioral neuroscience hail from a variety of different research traditions and theoretical backgrounds (e.g., animal psychology, neurophysiology) and have different technical expertise (e.g., expertise in assessing animal behavior or skill using *in vivo* circuit techniques). Given such differences, they do not necessarily agree about how to define terms typically used to designate cognitive functions (e.g., attention, working memory and motivation) and each field "contributes a distinctive vocabulary of terms and acronyms, all embedded to some degree or another in zeitgeists and conceptual frameworks" (Roediger, Dudai, and Fitzpatrick 2007: 1).

Although we do not encounter high-level theories in rodent behavioral neuroscience, researchers do have background assumptions about phenomena of interest, assumptions about what kinds of apparatus and tools are appropriate for addressing their empirical questions, theoretical understandings that inform the development of cognitive tasks and the use of intervention techniques as well as their understanding of how the tasks and tools they use actually work. Yet, differences in theoretical backgrounds, training and technical expertise across the field correlate with differences across researchers with respect to all of the different kinds of "ideas" that Hacking itemizes.

We encounter similar diversity with respect to Hacking's category of "things"; a number of different tasks may be used to study the "same" cognitive function, and not only do investigators differ with respect to what they regard as the most appropriate task or apparatus, but even when they use the same tool to investigate the same function, it is not uncommon for them to vary overall features of the task (e.g., stimuli, intertrial intervals) slightly (e.g. Sullivan 2009). Researchers also have different intuitions with respect to which tasks are most *appropriate* for measuring which functions and are granted the freedom to use those tasks and task parameters they deem most suitable for achieving their investigative aims, just so long as they provide good reasons for their choices from the perspective of peer review.

Differences in training also may impact the design and implementation of rodent behavioral experiments. For example, an expert in rodent behavior may be privy to aspects of an experimental design that may impact the behavioral performance of a mouse in a cognitive task (e.g., over-handling of the animal during different phases of the experiment) and confound the establishment of causal relationships between neural activity and behavior. They thus may modify aspects of the experimental protocol or specific task parameters with the aim of eliminating these confounds. In contrast, a researcher who is an expert in using neurophysiological techniques may be concerned with a different set of potential confounds having to do with consequences downstream of a pharmacological intervention. Such potential differences in epistemic standards that correlate with differences in expertise may thus exist. However, it is widely recognized that such methodological differences may result in differences in findings across laboratories purportedly investigating the same cognitive function (See for example, Crabbe, Wahlsten, Dudek 1999; Graybeal, Bachu, Mozhui *et al.* 2014; Sullivan 2009). This means that findings from multiple different research studies purportedly investigating mechanisms of the same phenomena cannot readily be integrated into unified explanations of common phenomena. And yet, discovering the neural mechanisms of cognition is not something that can take place in a single lab or in the context of a single research study. It requires contributions from many laboratories, not only to produce piecemeal findings about components of the neural mechanisms that give rise to a given cognitive function, but also to reproduce findings across laboratories (Beraldo, Palmer, Memar *et al.* 2019; Button, Ioannidis, Mokrysz, *et al.* 2013).

With respect to Hacking's category of "marks", researchers working in different laboratories also may use a variety of different tools for collecting, analyzing, and interpreting data, and employ different strategies to probe for and reduce error. Choices about which data analysis tools to use, what kinds of errors to probe and control for also vary with respect to one's training and technical expertise. An additional issue is that experiments in rodent cognitive neuroscience combine tools for assessing cognition with state-of-the-art visualization and/or intervention technologies. Yet, the error characteristics, especially of newer intervention and visualization technologies (e.g., optogenetics (Sullivan 2018), may not yet be known. A final and related issue is the lack of emphasis on the development of behavioral experiments that carefully individuate psychological functions involved in task performance and insure that the criterion of construct validity – that a given cognitive task actually measures the cognitive function it is intended to measure – is met prior to moving to experiments directed at identifying the neural underpinnings of these

functions (e.g. Krakauer, Ghazanfar, Gomez-Marin, *et al.* 2016; Niv 2020). There are thus epistemic blind spots in rodent behavioral neuroscience that are obstacles to the field advancing an understanding of the neural underpinnings of psychological functions.

Given the aforementioned observations, there is no sense in which the relationship between "ideas, materials, marks and [the] manipulation of marks" that we encounter in contemporary rodent behavioral neuroscience is stable, nor is the field on a trajectory towards stability. Yet, instability of the kind we find here is regarded by some neuroscientists themselves (i.e., those that I have cited in this section) as a barrier to progress in their field. Particularly in translational areas of cognitive neuroscience, in which the aim is to develop effective therapeutic interventions to treat neuropsychiatric and neurodegenerative disease-related cognitive impairments, the importance of reproducibility and the gradual coordinated accumulation of stable knowledge is regarded as essential for progress. In recent years, large-scale and smaller-scale collaborative grass roots initiatives have emerged with an eye towards stability of the kind Hacking describes. I turn now to analysis of these initiatives.

5. *Recent developments in rodent behavioral neuroscience through Hacking's lens*

In the first two decades of the 21st century, several large-scale initiatives were established in order to accelerate the discovery of novel therapeutic interventions to treat cognitive impairments in neuropsychiatric and neurodegenerative disease. Representative examples include the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS) initiative (e.g. Carter and Barch 2007; Moore, Geyer, *et al.* 2013), NEWMeds (e.g. Stensbøl and Kapur 2015), and the US National Institute of Mental Health's Research Domain Criteria Project (NIMH RDoC) (e.g., Insel, Cuthbert, Garvey *et al.*, 2010; Cuthbert & Kozack 2013). Each of these initiatives have brought together rodent behavioral neuroscientists, clinical researchers, cognitive neuroscientists working with humans and/or animal models, systems neuroscientists and members of the pharmaceutical industry with the aims of (1) developing more representative mouse models of neurodegenerative and neuropsychiatric diseases ("things"), (2) identifying a set of collaboratively agreed-upon psychological constructs corresponding to functions regarded as impaired in these diseases, ("ideas") (3) improving tools for assessing cognition in humans and mice ("things" and "manipulation of marks"), and (4) increasing the similarity of tools used for the behavioral assessment of cognitive functions across researchers and species ("things").

One way to understand the aims of these initiatives is to develop *stable* knowledge about the neural underpinnings of cognition and disruptions in neural circuitry that underlie these impairments in order to identify those circuits that may be targeted for therapeutic intervention. The measures that researchers involved in these initiatives regard as essential to these goals, are to develop a shared set of theoretical constructs (e.g., cognitive control, working memory) and types of apparatus/tasks (e.g., the Jitter orientation visual integration task (JOVI)) that are to be standardized across researchers working with human subjects and animal models, as well as a shared set of materials (e.g., apparatus, tasks, mouse models of disease) that are to be used in the drive to identify novel targets for therapeutic intervention. As Hacking claims, data interpretation relies on an investigator's background assumptions and theoretical commitments. Insofar as investigators involved in these initiatives are committed to a discreet set of theoretical constructs and general definitions of those constructs, the hope is that there will be some degree of consensus in how to interpret the data arising out of human and animal research. Thus, these initiatives are at least in theory aiming for coherence among Hacking-like elements – concepts, materials, and marks – that are disunified in cognitive neuroscience more generally.

These large-scale government supported research initiatives are on-going, however, to date, they have not produced stable knowledge or major advances in our understanding of cognition and cognitive dysfunction. While a number of reasons may be cited – clearly this is research on the cutting-edge and it is still early days – but one feature that such initiatives lack is an infrastructure to facilitate the stabilization of "ideas, things, and marks" across research groups and laboratories. It is one thing to point to changes that need to be made to experimental practice to facilitate progress and the production of stable knowledge and another thing for researchers to collaboratively implement these stabilization strategies across laboratories to achieve these goals.

The recent development of more grass-roots collaborative initiatives in rodent behavioral neuroscience (e.g., Beraldo, Palmer, Memar, *et al.* 2019; Dumont, Salewski, Beraldo, *et al.* 2020; Sullivan *et al.* 2020) and systems and computational neuroscience (with a focus on rodent behavioral research) (e.g., International Brain Laboratory 2017; Wool 2020) to accelerate discovery in these fields is suggestive that some researchers believe that achieving stability with respect to "ideas, things and marks" requires an unprecedented level of coordination across labs and research groups and an infrastructure similar to that found in other areas of science that have achieved stability historically, including physics and genomics (International Brain Laboratory 2017; Beraldo, Palmer, Memar *et al.* 2019). My aim in the rest of this section is to briefly evaluate these two grass-roots initiatives through Hacking's lens.

The first such initiative I want to consider has emerged around a novel platform, the *Mouse Translational Research Accelerator Platform (MouseTRAP)* (Sullivan *et al.* 2020). Spearheaded by researchers at Western University, MouseTRAP is centered on a touchscreen cognitive testing system for rodents, the Bussey-Saksida touchscreen system (e.g., Bussey, Muir, Robbins 1994; Bussey, Holmes, Lyon, *et al.* 2012; Bussey, Rothblat, Saksida 2001). The system consists of an operant chamber with a touchscreen upon which visual stimuli are presented. Rodents are trained and tested on different cognitive tasks using these visual stimuli and are required to respond directly to the visual stimuli with nose-pokes. Correct choices are rewarded with a drop of strawberry milkshake or a food pellet. There are currently over 20 different rodent touchscreen-based tasks for assessing cognitive functions in rodents ranging from working memory to cognitive flexibility to decision-making. The tasks are fully automated, ensuring the accuracy of task parameters and measures, and infrared beams and video tracking devices are used to monitor an animal's behavior while it performs in the apparatus. These features make the testing system and associated tasks readily standardizable across laboratories, allowing researchers all over the globe to use the same apparatus, stimuli, task parameters, appetitive rewards and data production and data analysis techniques.

In order to increase the reproducibility of rodent behavioral research and in response to increasing demand for the technology, the Bussey-Saksida touchscreen system was commercialized in 2009. Bussey, Saksida and colleagues published three invited papers in *Nature Protocols* (e.g., Horner, Heath, Hvoslef-Eide, *et al.*) with step-by-step instructions on how to prepare animals for training in the apparatus, how to pretrain and train the animals and how to analyze the behavioral data. As of December 2020, over 300 different research groups in more than 200 research institutes in at least 26 countries are using the touchscreen technology (Dumont, Salewski, Beraldo 2020). In 2018, two novel Open Science platforms were established to facilitate pre-publication knowledge-sharing (touchscreencognition.org) and data-sharing (mousebytes.ca) among members of the rodent touchscreen community. A primary aim of these Open Science platforms is "to create a community of scientists who share common methodology and are united in the goals of increasing methodological transparency and improving the reliability and reproducibility of research findings" (Sullivan, Dumont, Memar *et al.* 2020: 10).

If we consider MouseTRAP from the vantage point of Hacking's taxonomy of elements of experimental science, it possesses those features that lend themselves to the development of stable science – efforts are in fact being made to ensure that researchers share a common methodology for conducting research into the neural underpinnings of cognition, that they share ideas – for exam-

ple, empirical questions directed at specific phenomena (e.g., cognitive functions and impairments), topical hypotheses that relate specific understandings of those phenomena to what is observed in the laboratory, an understanding of how the apparatus works in the collection and production of data. They also share "things" in common – the targets of investigation (e.g., rodents, mouse models of disease), how to prepare those targets (as specified in the published protocols, and standardized operating procedures that are available on touchscreencognition.org), the touchscreen operant chamber itself and the tools (e.g., video-tracking devices) used to collect data. Those researchers who elect to use the methodology also share "marks and the manipulation of marks" – techniques of data assessment and analysis in common, and they are also at liberty to take advantage of Open Science platforms that allow them to share their knowledge, input and visualize their data and integrate and compare their data with data from other laboratories using the same methodology. MouseTRAP is suggestive of the fundamental role that scientists themselves must play to collaboratively produce stable science as Hacking conceives of it.[4]

Another notable collaborative grass-roots initiative is the International Brain Laboratory (IBL 2017; Wool 2020). It consists of ~80 researchers from 22 experimental and theoretical laboratories across the globe who are collaboratively aiming to identify the neural basis of decision-making. These researchers are using a standardized "steering-wheel task for head-fixed mice" in order to identify those brain areas that are involved in "decisions" made on the basis of "visual perception" and "history of reward" (IBL 2017: 1213). Using the same behavioral task across 22 laboratories, researchers in each laboratory will "record from many different brain areas" during task performance "using multiple recording modalities to build up a dense dataset of activity measurements during the task" (IBL 2017: 1213). These datasets will then be analyzed using computational techniques in order to understand how multiple brain regions interact during this task. IBL was developed because of the observed success of "team science" in other areas of science including physics and genomics. Moreover, "a critical IBL mandate is to ensure that theory and experiment converge at the ground level, and perpetually throughout [the] scientific process" (Wool, International Brain Laboratory 2020: 105).

IBL emphasizes the importance of bringing Hacking's elements of stability into a kind of consilience. The community seeks to ensure that members share theoretical and background knowledge, the same physical materials and

[4] I have referred to such collaboration as "coordinated pluralism" (2018). Knorr Cetina's (1999) concept of "epistemic culture" and Ankeny and Leonelli's concept of "repertoire", also may be used to shed interesting and important light on how stability or stable knowledge are collaboratively achieved in science.

tools and the same data production and data analysis techniques. They even emphasize the importance of "stabiliz[ing] large-scale collaborative science in traditional academia" in order to achieve the goal of "understanding the neural computations that support decision-making" (IBL 2018: 1213).

One way to conceive of these grass-roots initiatives is that they regard the accumulation of knowledge of the neural underpinnings of cognition to require what I described in Section 3 as "productive disunity". Such disunity involves the collaborative breaking off of smaller groups of investigators from how practice in a given area of science is traditionally done, in instances in which sticking with tradition involves "counterproductive disunity" that is antithetical to progress. It is an interesting question whether laboratory sciences like those Hacking (1992) uses as a basis for understanding stability began with small-scale collaborative revolutions much like these ones.

6. *Conclusion*

I want to end by teasing out some implications of my analysis for the unity of neuroscience and say something briefly, from the perspective of Hacking's lens, about the kinds we are liable to encounter in rodent behavioral neuroscience if such grass-roots initiatives are successful.

First, it is relevant to note that the experimental apparatuses at the heart of both of these initiatives satisfy Hacking's condition that laboratory sciences "create new phenomena". Nowhere in the world (as far as I know), except in laboratories that use rodent operant touchscreens, do we encounter rodents interacting with and engaging in cognitive tasks with computer touchscreens. Similarly, we do not encounter head-fixed mice out and about in the world turning steering wheels in response to visual stimuli. The kinds of cognitive functions under study using these apparatuses are created in laboratories. This does not make them any less real, but it is important to recognize the precise type of workmanship that goes into creating them (e.g. Boyd 2000). Moreover, if these small-scale initiatives are ultimately successful, they may yield what might be dubbed "coordinated kinds" (Mattu and Sullivan in press) – the result of the concerted alignment of conceptual and methodological practices across discrete research groups with respect to "ideas, things, and marks". To the extent that different such research groups emerge in cognitive neuroscience and are successful, organizing their practices around discrete sets of concepts, apparatus, tools, and data, we might imagine a plurality of discrete taxonomies of cognitive kinds that are stable but isolated from each other – a kind of "promiscuous realism" (Dupre 1993).[5]

[5] Thanks to Muhammad Ali Khalidi for this characterization – an idea to be worked out on an-

Second, insofar as the creation of these phenomena and investigations into their mechanisms are to be collaboratively subserved by small groups of researchers, and if such collaborations are successful in bringing about a kind of local stability – the kind of findings such research groups make about neural mechanisms are likely to be domain-specific – specific, for example, to those "ideas, things, and marks" that these groups collaboratively bring into consilience to achieve stable knowledge. This is consistent with Hacking's (1991) idea that successful stability is consistent with disunity – that it actually requires disunity – it requires a kind of isolation of a domain from factors that are antithetical to its stability.

On a final note, Hacking would likely be skeptical that these collaborative initiatives, even if they can yield stable knowledge about the neural mechanisms of cognition in rodents, will ultimately shed light on the mechanisms of human cognition, because "human kinds" are "unstable" in ways that make them unamenable to experimental control (e.g. Hacking 1995; 1999). Partially for reasons of space, evaluating and responding to such skeptical concerns will have to be saved for another occasion.

Jacqueline Sullivan
jsulli29@uwo.ca
University of Western Ontario

References

Ankeny, Rachel, Sabina Leonelli, 2016, "Repertoires: A post-Kuhnian perspective on scientific change and collaborative research", in *Studies in History and Philosophy of Science*, 60: 18-28.

Beraldo, Flavio, Daniel Palmer, Sara Memar, *et al.*, 2019, "MouseBytes. An open-access high-throughput pipeline and database for rodent touchscreen-based cognitive assessment", *eLife*, 8: e49630.

Bekinschtein, Pedro, Brianne Kent, Charlotte Oomen *et al.*, 2013, "BDNF in the Dentate Gyrus is required for consolidation of "Pattern-Separated" memories", in *Cell Reports*, 5: 759-768.

Bechtel, William, 2008, *Mental Mechanism: Philosophical Perspectives on Cognitive* Neuroscience, Routledge, New York.

Boyd, Richard, 2000, "Kinds as the workmanship of men: realism, constructivism, and natural kinds", in Julian Nida-Rümelin, ed., *Rationalität, Realismus, Revision*, De Gruyter, Berlin: 52-89.

Bussey, Timothy, Andrew Holmes, Louisa Lyon, *et al.*, 2012, "New translational assays

other occasion.

for preclinical modelling of cognition in schizophrenia: The touchscreen testing method for mice and rats", in *Neuropharmacology*, 62, 3: 1191-1203.

Bussey, Timothy, Janice Muir, Tim Robbins, 1994, "A novel automated touchscreen procedure for assessing learning in the rat using computer graphic stimuli", in *Neuroscience Research Communications*, 15, 2: 103-110.

Bussey, Timothy, Lisa Saksida, Lawrence Rothblat, 2001, "Discrimination of computer-graphic stimuli by mice: A method for the behavioral characterization of transgenic and gene-knockout models", in *Behavioral Neuroscience*, 115, 4: 957-960.

Button, Katherine, John Ioannidis, Claire Mokrysz, *et al.*, 2013, "Power failure: Why small sample size undermines the reliability of neuroscience", in *Nature Reviews Neuroscience*, 14: 365-376.

Carter, Cameron, Deanna Barch, 2007, "Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: the CNTRICS initiative", in *Schizophrenia Bulletin*, 33, 5: 1131-1137.

Craver, Carl, 2007, *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, Oxford University Press, Oxford.

Crabbe, John, Douglas Wahlsten, Bruce Dudek, 1999, "Genetics of mouse behavior: interactions with laboratory environment", in *Science*, 284: 1670-1672.

Cuthbert, Bruce, Michael Kozack, 2013, "Constructing constructs for psychopathology: the NIMH research domain criteria", *Journal of Abnormal Psychology* 122(3): 928-937.

Duhem, Pierre, [1906] 1954, *The Aim and Structure of Physical Theory*, Princeton University Press, Princeton.

Dumont, Julie, Ryan Salewski, Flavio Beraldo, 2020, "Critical mass: the rise of a touchscreen technology community for rodent cognitive testing", in *Genes, Brain and Behavior*: e12650.

Dupré, John, 1993, *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Harvard University Press, Cambridge MA.

Graybeal, Carolyn, Minusa Bachu, Khyobeni Mozhui, *et al.*, 2014, "Strains and stressors: An analysis of touchscreen learning in genetically diverse mouse strains", in *PLOS ONE* 9: e87745.

Hacking, Ian, 1999, "Making up people", in Mario Biagioli, ed., *The Science Studies Reader*, Routledge, New York: 161-171.

Hacking, Ian, 1995, "The looping effects of human kinds", in Daniel Sberber *et al.*, eds., *Causal Cognition: A Multidisciplinary Debate*, Clarendon Press, Oxford: 351-383.

Hacking, Ian, 1992, "The self-vindication of the laboratory sciences", in Andrew Pickering, ed., *Science as Practice and Culture*, University of Chicago Press, Chicago: 29-64.

Hacking, Ian, 1991, "Disunified sciences", in R.J. Elvee, ed., *The End of Science?*, University of America Press, Lanham.

Hacking, Ian, 1983, *Representing and Intervening*, Cambridge University Press, Cambridge.

Horner, Alexa, Christopher Heath, Martha Hvoslef-Eide, *et al.*, "The touchscreen operant platform for testing learning and memory in rats and mice", in *Nature Protocols*, 8, 10: 1961-1984.

Insel, Thomas, Bruce Cuthbert, Marjorie Garvey, *et al.*, 2010, "Research domain criteria (rdoc): toward a new classification framework for research on mental disorders", *American Journal of Psychiatry*, 167, 7: 748-751.

International Brain Laboratory, 2017, "An international laboratory for systems and computational neuroscience", *Neuron*, 96, 6: 1213-1218.

Knorr Cetina, Karin, 1999, *Epistemic Cultures: How Science Makes Knowledge*, Harvard University Press, Cambridge MA.

Kuhn, Thomas, 1962, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Kraukauer, John, Asif Ghazanfar, Alex Gomez-Marin, *et al.*, 2017, "Neuroscience needs behavior: correcting a reductionist bias", in *Neuron*, 93, 3: 480-490.

Mattu, Jaipreet and Jacqueline Sullivan, (in press), "Classification, kinds, taxonomic stability and conceptual change", in *Agression and Violent Behavior.*

Mayo, Deborah, 1996, *Error and the Growth of Experimental Knowledge*, University of Chicago Press, Chicago.

Mayo, Deborah, 1991, "Novel evidence and severe tests", in *Philosophy of Science*, 58, 4: 523-552.

Moore, Holly, Marjorie Geyer, Cameron Carter, *et al.*, 2013, "Harnessing cognitive neuroscience to develop new treatments for improving cognition in schizophrenia: CNTRICS selected cognitive paradigms for animal models", in *Neuroscience Biobehavioral Reviews*, 9: Pt B:2087-2091.

Nagel, Ernest, 1961, *The Structure of Science: Problems in the Logic of Scientific Explanation*, Harcourt, Brace & World, New York.

Niv, Yael, 2020, "The primacy of behavioral research for understanding the brain", *PsyArXiv*: October 22; DOI: 10.31234/osf.io/y8mxe.

Oppenheim, Paul and Hilary Putnam, 1958, "Unity of science as a working hypothesis", in *Minnesota Studies in the Philosophy of Science*, 2: 3-36.

Piccinini, Gualtiero, and Carl Craver, 2011, "Integrating psychology and neuroscience: Functional analysis as mechanism sketches", in *Synthese*, 183, 3: 283-311.

Roediger III, Henry, Yadin Dudai, Susan Fitzpatrick, 2007, *Science of Memory: Concepts*, Oxford University Press, Oxford.

Schickore, Jutta, 2005, "Through thousands of errors we reach the truth – but how? On the epistemic roles of error in scientific practice", in *Studies in History and Philosophy of Science Part A*, 36, 3: 539-556.

Schikore, Jutta, 2020, "The structure and function of experimental control in the life sciences", in *Philosophy of Science*, 86, 2: 203-218.

Stensbøl T.B., Kapur S., 2015, NEWMEDS special issue commentary *Psychopharmacology* (Berl), 232(21-22): 3849-3851, DOI: 10.1007/s00213-015-4083-y

Sullivan, Jacqueline, Julie Dumont, Sara Memar, *et al*., 2020, "New frontiers in translational research: touchscreens, open science and the mouse translational research accelerator platform", in *Genes, Brain and Behavior*: e12705.

Sullivan, Jacqueline, 2018, "Optogenetics, pluralism and progress", in *Philosophy of Science*, 85, 5: 1090-1101.

Sullivan, Jacqueline, 2017, "Coordinated pluralism as a means to facilitate integrative taxonomies of cognition", in *Philosophical Explorations*, 20, 2: 129-145.

Sullivan, Jacqueline, 2016, "Construct stabilization and the unity of neuroscience", in *Philosophy of Science*, 83, 5: 662-673.

Sullivan, Jacqueline, 2014, "Stabilizing mental disorders: prospects and problems" in Harold Kincaid and Jacqueline Sullivan, eds., *Classifying Psychology: Mental Kinds and Natural Kinds*, MIT Press, Cambridge MA.

Sullivan, Jacqueline, 2009, "The Multiplicity of Experimental Protocols: A Challenge to Reductionist and Non-reductionist Models of the Unity of Neuroscience", in *Synthese*, 167: 511-539.

Suppes, Patrick, 1984, *Probabilistic Metaphysics*, Blackwell Publishers, Oxford.

Wool, Lauren, International Brain Laboratory, 2020, "Knowledge across networks: how to build a global neuroscience collaboration", in *Current Opinion in Neurobiology*, 65: 100-107.

Past Present

Le rationalisme appliqué

G. Bachelard

A dialogical philosophy: Bachelard's Introduction to *Le Rationalisme appliqué*

Lucie Fabry

In 1940, Gaston Bachelard became a Professor in Philosophy of Science at the Sorbonne and Director of the Institute of Science and Technology at the *rue du four* in Paris. He published *La Philosophie du non* at the Presses Universitaires de France (Bachelard 1940; Eng. tr. 1968a), which was his 13th book. To that date, his books had focused mostly on the history of physics and chemistry, even though Bachelard's interests extended far beyond these disciplines and included notably psychoanalysis (Bachelard 1938a; Eng. tr. 2002a; 1938b; Eng. tr. 1968b) and literary criticism (Bachelard 1932; Eng. tr. 2013; 1939; Eng. tr. 1986). Between 1940 and 1949, however, Bachelard did not publish any epistemological book. This nearly 10-year gap in Bachelard's prolific epistemological production is explained by the Second World War, a period during which he continued to write intensively but focusing on the powers of imagination (Bachelard, 1942; Eng. tr. 1983; 1943; Eng. tr. 1988; 1948a; Eng. tr. 2002b; 1948b; Eng. tr. 2011). As a counterpart to his works on the psychology of the scientific mind, he thus dedicated himself to the psychology of aesthetic emotions, inquiring how poetical images reverberate in us. The publication of *Le Rationalisme appliqué* in 1949 marked Bachelard's return to epistemological concerns.

The following text, available for the first time in an English translation by Gennaro Andrea Lauro, is a preliminary version of what would become the introduction to *Le Rationalisme appliqué* (Bachelard 1949). Bachelard published that text separately in February 1947 in the first issue of the international journal *Dialectica* (Bachelard 1947). The version of 1947 and that of 1949 are almost identical, except for minor revisions. We can, therefore, regard its translation as a first step towards making *Le Rationalisme appliqué* accessible to the English readers, hoping that it may lead to a translation of the entire work. It is nonetheless interesting to wonder why Bachelard chose to publish that introduction separately in *Dialectica*, and I will start by questioning the meaning and function this text had in the context of its first publication.

philinq IX, 1-2021, pp. 211-230
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.362

1. *Dialectica: International review of philosophy of knowledge*

Bachelard was one of the founding members of *Dialectica: International review of philosophy of knowledge*, along with Paul Bernays, a Swiss mathematician close to David Hilbert, and Ferdinand Gonseth, a Swiss mathematician and philosopher with whom Bachelard had strong intellectual affinities. The three authors opened the first issue of their journal with an Editorial published in English, French and German, which stated how strongly the Second World War had shaken the faith in the value of scientific knowledge: "Some people say: we refuse to accept this century in which we live; we refuse its knowledge and its inventions; we refuse to accept this civilization for which we feel ourselves less and less conjointly responsible, its technical ideals, its false values." (Bachelard *et al.*, 1947: 1). Despite their understanding of what motivates such rejection of the "scientific era", the authors refused that search for a way out and instead claimed that "the remedy against the dangers of knowledge is better knowledge" (1). They presented their international journal as part of "a sustained philosophical effort" which was required in order to keep Science "on the horizons of human values" (1). An indication that *Dialectica*'s reflections on science had a civilizational scope is also found in the text that concludes the first issue: the French translation of a speech by Pope Pius XII to the members of the International Congress of Philosophy held in Rome in November 1946 (Pius XII 1947). In this speech, the Pope affirmed that both (Catholic) religion and philosophy stood against "a certain pessimistic irrationalism" (Pius XII 1947: 110; Eng. tr. is our own) and wanted to lead the new generation towards the "sublimation of human tendencies in favor of superior ideals" (109), in spite of "the extreme atrocities that this youth has had to endure in recent years" (109). It would be excessive to claim that the choice to publish that allocution implied that the funders of the review subscribed entirely to the Pope's view – the journal also had the ambition to keep track of the activities that took place in and around these international congresses. This choice of publication may, however, indicate that the authors aspired to a broad syncretism around the belief that a better understanding of scientific thought could help us achieve a better humanity.

Bachelard's article was published in the pages which directly followed the Editorial: in that sense, it completed the journal's Manifesto and specified what conception of science was able to support these humanistic values. In Bachelard's article, however, the question of the dangers of technology and the quest for the norms and values of scientific civilization is almost entirely set aside to focus on a more traditional epistemological question: what is science, and how can it achieve true knowledge of reality? *Dialectica* as a whole did give

precedence to the analysis of scientific activity over the consideration of broad humanistic issues. The Journal would soon welcome publications from the most prominent scientists of that time, with a special issue in November 1948 on the notion of complementarity in atomic physics edited by Wolfgang Pauli (1948) including articles from Niels Bohr (1948), Albert Einstein (1948), Louis de Broglie (1948), Werner Heisenberg (1948) or Hans Reichenbach (1948); in December 1958, a remarkable double issue on mathematics and formal systems included, among others, texts from Rudolf Carnap (1958), Kurt Gödel (1958), Reuben Goodstein (1958) and Thoralf Skolem (1958). The founders of the journal refused, however, to separate these works on science from a broader philosophical framework, as indicate the very name *Dialectica* and the fact that the first issue was dedicated to that notion of dialectics, which was deliberately chosen as a broad and traditional philosophical concept. Even though this journal constitutes in its current form the official organ of the European Society for Analytic Philosophy, it is interesting to keep in mind that its original editorial line knew no divide between continental and analytic philosophy: these articles from Hans Reichenbach, Rudolf Carnap, Karl Popper (1978) or Alfred J. Ayer (1958) were peacefully cohabiting with texts from Jean Piaget (1950; 1954; 1959), Carl Gustav Jung (1951), Eugène Dupréel (1957) or Raymond Ruyer (1959).

2. *A philosophical dialogue between theoreticians and experimenters*

Let us now consider Bachelard's text a little more closely. At its core stands the diagram on page 235, of which Bachelard may have been the only one to believe that it is so clear that "it is barely necessary to comment on it" (234). This diagram displays philosophical tendencies: at the center stands a duo, "applied rationalism and technical materialism". Other trends stand on a ladder that goes up and down that center: Formalism and Conventionalism aligned with the ascent that leads to Idealism; Positivism and Empiricism aligned with the descent that leads to Realism. The aim of the remaining parts of this presentation will be, simply and solely, to understand what Bachelard meant when he drew that diagram. I will start by focusing on the center, "applied rationalism and technical materialism", which designate, as a couple, Bachelard's own position. I will then specify the role Bachelard gave to the other philosophical tendencies as he displayed them around that center.

One can be surprised by the fact that Bachelard did not give one name, but two complementary ones to his philosophy of science: "applied rationalism and technical materialism". He also designated it earlier in the text as "an *applied rationalism* and an *instructed materialism*" (234). Not only did he use a

combination of two expressions, but each of these expressions has itself a dualistic character: the name of a philosophical current is always associated with an adjective that specifies it. Jean-Claude Pariente (2015: 251; Eng. tr. is ours) remarked that "these expressions are always syntactically dualist and semantically auto-correcting. These expressions are composed of a noun and an adjective. The function of the adjective is to straighten, fix the effect of the noun and warn against its heaviness. They must always be taken with an implicit 'but' in it." The combination of two of such expressions looks like a series of checks and balances: Bachelard's philosophy is rationalism *but* it is applied *but it is also* materialism, *but* it is technical and instructed.

That way of naming his philosophy reflects Bachelard's commitment to a "dialogical philosophy" (*philosophie dialoguée*), that is, a philosophy that stems from a dialogue and is constituted by it. That notion can be regarded as a Socratic legacy, and so does Bachelard's use of the notion of dialectics in a sense that is close to its Greek etymology, referring to a dialogue. Less traditional, however, are the protagonists of that dialogue, and what they are talking about. The philosophical dialogue Bachelard referred to at the beginning of "La Philosophie dialoguée" was not, he insisted, a dialogue between philosophers, but a dialogue between scientists. Bachelard claimed that philosophers do not know how to have a fruitful philosophical dialogue, only scientists do. By that, he did not mean that scientists should practice philosophy on their spare time, and would then become, thanks to their scientific knowledge, the best philosophers. Even though he claimed that a precise knowledge of science was required to do good philosophy, he also deplored that scientists, once they start philosophizing, often become philosophers like any others, prompt to commit the same mistakes (see for instance Bachelard 1953: 19-20). Bachelard did not claim that scientists should turn to philosophy, but rather that scientific discussions *as such* constitute the best philosophical debates one can hope for. The dialogue he presented occurs, more specifically, between a specialist of theoretical physics and a specialist of experimental physics. Their conversations constitute the concrete and daily relationship between theory and experience, which, according to Bachelard, is the place to look to seriously address the question of the relationship between mind and reality. However, why should we consider their discussions as a philosophical dialogue, considering that they are not talking about philosophy but rather exchanging precise theoretical or experimental information? According to Bachelard, the theoretician and the experimenter adopt, by virtue of their professional specializations and without necessarily being aware of it, two different philosophies: rationalism and empiricism. Rationalism, as Bachelard understands it here, gives primacy to theory: it is above all concerned with the search for theoretical coherence,

completion and rigor. It defines, from that perspective, the role of experiment, claiming that theoretical hypotheses should lead experimental programs and that experimental knowledge has little value without its theoretical interpretation. Empiricism, on the other hand, gives the primacy to experimentation: it is mostly concerned by experimental precision and claims that we should consider the facts even when they do not match our theories, arguing that facts can suggest new theories, lead to a modification of admitted theories or even to their rejection. In philosophy, according to Bachelard, these attitudes towards theory and experimentation have crystallized in two antithetic views on knowledge and science. Therefore, when a debate occurs between rationalists and empiricists, there is little hope that it will lead to a reduction of their antagonism. In the daily practice of physics, however, rationalism and empiricism exist as two professional attitudes: theoreticians are professional rationalists, experimentators professional empiricists. In this case, these attitudes are shared between people that know they have to work together. Despite the growing division of labor and specialization which allowed that theoretical and experimental physics are practiced by different persons with different trainings and professional habits, it is nonetheless clear, according to Bachelard, that physics as a discipline lies at their intersection: "If one of the two terms is missing, we can still do experiments and we can still do mathematics, but we cannot participate in the scientific activity of contemporary physical science" (234). This solidarity between theoretical and experimental physics takes very concrete forms: for instance, "no physicist would spend "his credit" to build an instrument with no theoretical destination" (233).

Bachelard's ambition as a philosopher was to draw lessons from the fact that rationalism and empiricism are actually combined in scientific activities, as two complementary professional attitudes. This is what prevented him from adopting what he calls "monodromic philosophies", (*philosophies monodromes*; Bachelard 1949: 159): pure rationalism or empiricism, and led him to look for a way to combine them. Dualist expressions such as *applied rationalism* designate that philosophical combination. The expression *applied rationalism* was built in analogy with the distinction between *pure* and *applied mathematics.* However, this analogy may be misleading if one considers that pure mathematics does exist as an autonomous scientific field: even if the possibility of finding long-term applications is one of the reasons why this field receives funding, it is a fact that not all mathematics is applied mathematics. Following that analogy, one could assume that pure and applied rationalism could coexist, equal in value, investigating different aspects of rationality. However, Bachelard focused on physics, where mathematics must, by definition, be applied to the knowledge of physical reality. He precisely intended to elucidate the conditions

under which mathematics can be made to serve such purpose and stated that they take the form of an endless dialogue between theory and experiment. From that perspective, pure mathematics can be nothing but flawed physical theories, that is, theories that tend to withdraw from experimental control or are unable to suggest new experiments. By analogy, Bachelard argued that pure rationalism misunderstood the way the mind constitutes knowledge because it underestimated the way our reasoning is shaped by experience. He addressed a symmetrical accusation to empiricism or materialism – it seems that Bachelard regarded these terms as synonymous in that text, even though a later text, *Le Matérialisme rationnel* (Bachelard, 1953), gave another, more specific sense to the notion of materialism. As a result, were Bachelard to make his own the notion of materialism, it would only be with the qualification that materialism is technical or instructed, which are two manners of underlining the way experience is shaped by theory.

That ambition to apprehend the solidarity between theory and experiment led Bachelard to be conceptually innovative. Besides the expressions of *applied rationalism* and *instructed materialism*, he elaborated the notion of *phenomenotechnique*, which can be regarded as one of Bachelard's most significant contribution to the philosophy of science (Bontems 2010; Castelão-Lawless 1995; Chimisso 2008; Donatiello *et al.* 2018; Fabry 2019; Granger 1987; Gaukroger 1976; Rheinberger 2005). When he elaborated it in the early 1930s, and as he used it ever since, that notion is inseparable from another neologism, *noumenology* (Lamy, 2005). Noumenology is the name Bachelard gave to theoretical activities *considered in their relation to experiment*; phenomenotechnique is the name he gave to experimental activities *considered in their relation to theory*. The notion of phenomenotechnique insists on the technological character of a scientific experiment: it is not the mere observation of a fact but the production of a phenomenon. Even though they serve different ends—when industry and science are not actually combined—a laboratory and a factory are analogous, according to Bachelard, in the sense that both create artefacts that meet specific ends and comply to specific constraints. However, Bachelard insisted on the technological character of a scientific experiment only since it was indicative of its dependence towards theory. Indeed, he retained from Pierre Duhem (2006; Eng. tr. 1991) that "instruments are theories materialized", in the sense that the conception and use of scientific instruments rely on a set of theoretical assumptions. Hence, artificially produced phenomena "bear the stamp of theory throughout" (Bachelard, 2020; Eng. tr. 1984: 13). On the side of theory, Bachelard's concept of noumenology acknowledged that mathematical theories possess their own dynamic and may legitimately explore paths that do not seem to have any em-

pirical correlate. However, the notion of noumenology nonetheless insisted on the fact that such mathematical excursions would only be valued, in physics, according to their capacity to lead the successful production of phenomena that conform to its expectations. Even though the experiment relies on theoretical assumptions, as the notion of phenomenotechnique indicates it, the experiment may nonetheless address objections to these theoretical premises by departing from what was expected. Bachelard's dialectics between theory and experiment, or, more precisely, between noumenology and phenomenotechnique, thus designates the dialogue through which theory and experiment conjointly evolve and shape each other, as theoretical suggestions meet experimental objections, or unexpected experimental results look for their theoretical interpretation.

3. *Dialectics, doublets and idoneism*

In "La philosophie dialoguée", Bachelard assimilated his own concept of dialectics with two other conceptual innovations: Léon Brunschvicg's *doublets*, and Ferdinand Gonseth's *idoneism*. In an earlier text (Bachelard 1945), Bachelard had presented Léon Brunschvicg as his philosophical Master and claimed that *L'Expérience humaine et la causalité physique* (Brunschvicg 1922) offered an eminent analysis of the way theory relates to the experiment. He noticed Brunschvicg frequent use of distinctions such as *numbering number* and *numbered number*, *spacializing space* and *spacialized spaced*, and called such expressions "Brunschvicgian doublets", claiming they were built in reference to Spinoza's distinction between *natura naturans* (naturing nature) and *natura naturata* (natured nature) (239). Even though the actual intent of Spinoza (2020) when he borrowed this distinction to Scholasticism is subject to much discussion (Ramond 2011), it seems that Bachelard regarded it as a way of distinguishing God from its creation whilst claiming that they constitute the same thing, considered from two different perspectives. Similarly, Brunschvicg's doublets would distinguish two complementary ways of conceiving scientific objects whilst claiming that they cannot, in fact, be dissociated. The solidarity between theory and experiment implies, according to Bachelard's reading of Brunschvicg, that the same facts can be translated in "two languages": "the translation of scientific experiment and the translation of rational coherence" (Bachelard, 1945: 81; Eng. tr. is our own). If one looks back to Brunschvicg's work, however, it appears that Bachelard's reading minimized some aspects that made this work closer to pure rationalism, in the sense stated above, than applied rationalism. When Brunschvicg introduced the distinction between *numbering numbers* and *numbered numbers*, in the sense of a distinction be-

tween the activity of counting and the things counted, he insisted on the fact that the understanding of the prevalence of numbering numbers was a necessary condition for the emergence of mathematics and the further development of human knowledge (Brunschvicg 1922: 473-477). He significantly did not refer to Spinoza but to Malebranche (see for instance Robinet 1961 for an analysis of Malebranche's use of the notion of numbering number). The main thesis of his "philosophy of judgment" (Brunschvicg 1922: 474) was that science progresses as the mind gets more aware of its constitutive activity, instead of conceiving the world as a pre-existing given. One can, therefore, argue that even though Bachelard was deeply influenced by Brunschvicg's rationalism (Vinti 1997), he nonetheless departed from it progressively by giving a more important role to scientific experiment (Dagognet 1965).

The second author with whom Bachelard claimed to have affinities, Ferdinand Gonseth, seems to stand closer to his applied rationalism. Gonseth's notion of *idoneism* may be regarded as an attempt to express the way in which knowledge and its objects conjointly evolve in an open process of mutual adjustments (Bertholet, 1968; Bontems, 2013, 2018). Gonseth wanted to substitute this neologism – built from the French adjective *idoine* (which could be translated as adequate, appropriate to something) – to the traditional notions of *correspondence* or *adequacy* between knowledge and reality, because he considered that these notions conveyed the idea that knowledge and thought may exist as separate, static entities (Gonseth 1936; 1939). Instead, idoneism designates an adequation which is always in the making and perpetually transforms the entities that aspire to be adequate to one another. Interestingly, whereas Bachelard's main reference when he theorized his *applied rationalism* was mathematical physics, Gonseth elaborated his notion of *idoneism* by reflecting mainly on mathematics (Gonseth 1936). The philosophical problem of the relation between knowledge and reality could be addressed, Gonseth claimed, within the field of mathematics itself, when one considers, for instance, the axiomatization of mathematics. The process through which an axiomatic theory is built does imply, according to Gonseth, an evolving relationship between an abstract, theoretical object and a mathematical object which can be regarded as concrete and intuitive, even though it is not a physical entity. Since *The New Scientific Spirit*, Bachelard had similarly claimed that the dialectics between theory and experiment could, paradoxically, already be found in pure mathematics: "this need of application is felt just as strongly in pure mathematics, though there it is more hidden. It introduces an element of metaphysical duality into the mathematical sciences, which appear to be purely homogeneous […] every pure idea is accompanied by an imagined application, an example that does duty for reality […] In a reasonably clear-cut manner, mathematical

realism (in its various functional roles) sooner or later operates to *give body* to pure thought" (Bachelard 2020: 28; Eng. tr.: 4-5).

It is remarkable that both Bachelard and Gonseth, having a specific discipline dominantly in mind, intended to elaborate concepts that would apply to knowledge in general, or, at least, scientific knowledge. Bachelard claimed that his concepts, built in a reflection that focused mainly on physics, could also apply to pure mathematics. In contrast, Gonseth claimed that his idoneistic analysis of mathematics could also apply to the knowledge of the physical world. The strong similarities noticed between Bachelard and Gonseth theses seem, indeed, to indicate that such an extension of their concepts out of their initial field is possible. However, the philosopher and physicist Jean-Marc Lévy-Leblond (2017) has argued that Bachelard's dialectics between noumenology and phenomenotechnique may well be an accurate description of physics, but can hardly be applied to other scientific practices. In "La Philosophie dialoguée", Bachelard himself suggested that an analysis of "other sciences, such as mathematics, biology, sociology and psychology" (237) could lead to other kinds of philosophies. Throughout his epistemological work, the history of physics constituted Bachelard's favorite philosophical playground, even though he also offered in-depth analyses in on the history of mathematics (Alunni 2015) and chemistry (Bensaude-Vincent 2005; 2012). He regarded these sciences as different regions of knowledge, who have their own history and dynamics, and do not build a unified system (Bachelard 1949: 149 sq.). When questioning the history of these disciplines, he used the same set of epistemological concepts, wondering what forms of rationalism are applied in mathematics or chemistry. Concerning biology, sociology and psychology as sciences, Bachelard said very little. Does this indicate that these disciplines would have required a completely different epistemology? Since the 1960s, one of the leading questions among the historians and philosophers, who acknowledged some kind of Bachelardian legacy, has been the whether and how Bachelard's concepts could be useful to the study of biology or human and social sciences (Althusser 1965; Eng. tr. 2010; Bourdieu *et al.*, 1972; Eng. tr. 1991; Canguilhem 1965; Eng. tr. 2008; Granger 1967; Eng. tr. 1983; Rheinberger 2010).

4. *Two perspectives of weakened thoughts*

Let us go back to Bachelard's diagram on p. 235. The center, "applied rationalism and technical materialism", designates what he regarded as the legitimate way of conceiving the relation between theory and experiment, giving them an equal epistemological weight and stressing their deep solidarity. The other philosophies are ordered in what Bachelard called "two perspectives of

weakened thoughts which, on one hand, lead from rationalism to naive idealism and, on the other hand, from the technical materialism to naive realism" (235). These "*weakened* thoughts" are characterized by the fact that they have lost the balance between theory and experiment: once you start overestimating the importance of theory and considering it aside from experiment, you are on the path to idealism; if, on the contrary, you overestimate the importance of experiment alone, and consider it aside from theory, you are on the path to realism. To put it boldly, Bachelard's diagram classes the philosophical positions of his predecessors and contemporaries by indicating how they failed to be Bachelardian. It must be noticed, however, that his intent is not purely critical: he claimed that he would "mutilate philosophy of science" if he did not take into account that, aside from applied materialism and technical materialism, other philosophical attitudes are part of "the psychology of the scientific mind" (238). He agreed, at least for those philosophies that are not too far from the centre of his diagram, that they also play an active and positive role in scientific thought: that these are not philosophers' ratiocinations that have lost track of what science is, but philosophical attitudes of scientists, in a sense that I mentioned earlier. He also felt that the philosophies that are at the same rank in the centre of his diagram could somehow dialogue with each other. That is to say that formalism and positivism as professional ways of conceiving theory and experiment, respectively, could somehow be combined. Their combination would constitute a way of articulating theory and experiment that would be looser than the one that characterizes noumenology and phenomenotechnique, but that would nonetheless be quite satisfying.

What may make it difficult to understand Bachelard's diagram is that he had his own definitions of such notions as positivism, formalism, idealism or realism, definitions which he never plainly formulated but can be inferred from his use of these terms. All these *-isms* function like ideal types: it is less important for Bachelard to target specific philosophers than to identify broad philosophical attitudes. *Formalism* could be defined by the assumption that scientific theories are autonomous deductive systems, which stem from their own postulates and can be considered independently from the experiment, from the perspective of their logical consistency alone, even though conclusions can be drawn from these systems that we can submit to experimental control. *Positivism*, as Bachelard used that term, would stand for the assumption that scientific theories sum up our knowledge of facts and register regularities in observable phenomena, but do not pretend to reflect, accurately, what reality is. Formalism and Positivism are not incompatible: Bachelard did claim that a dialogue could be set between them. If we assume that a theory is a deductive system, which is built up independently of experience, we may

well admit that different theories could be applicable to the same set of facts, as they do not claim to reflect their inherent nature, but to provide an accurate representation of our data. We could look for philosophers who stood for such of a combination of formalism and positivism, perhaps among the tenants of logical empiricism. It is, however, more interesting to note the role that Bachelard gave to this philosophical position: he felt that such an alliance of positivism and formalism was unsatisfactory as a discourse on science in general but could nevertheless provide an accurate description of certain aspects of scientific practice. He was dissatisfied by formalism and positivism since they both state too loose a connection between theory and facts, in comparison to his concepts of noumenology and phenomenotechnique: formalism grants too much autonomy to mathematical reasoning by considering that the contact with experience is only terminal, instead of considering that it actually permeates conceptual activity throughout; positivism granted too much autonomy to experience by claiming that the role of a theory merely consists in summing up the facts, while Bachelard held that it actively produces them through phenomenotechnique. However, while positivism and formalism were deemed unfit for a depiction of these parts of contemporary physics which had reached the level of phenomenotechnique, Bachelard admitted that there were other parts of scientific activity where the connection between theory and experiment *actually was* looser. While asserting that his concepts of phenomenotechnique and noumenology provided a good understanding of the most advanced practices in contemporary physics and chemistry, he also acknowledged that these concepts made very specific demands on theory and experimentation. The belief that these demands were always met would lead to an idealized view of scientific practices: we should rather explore the diversity of theoretical and experimental practices that can be found in the past and present of these disciplines. Positivism could therefore describe correctly the status of emerging theories: in a subsequent work, Bachelard (1953) studied the history of chemistry and argued that this discipline was once positivist. Looking at chemistry in Kekulé's time, it could very well be said, for example, that the geometric representation of the four valencies of carbon was a theoretical synthesis of experimental knowledge that did not claim to represent what carbon really looked like at the microscopic level. However, the status of this representation evolved as the phenomenotechnique of mass spectrometry, as well as the theories of microphysics, developed and gave it a higher objective value: any positivist account of contemporary representations of carbon would therefore be condemned as reductive. Bachelard did not only claim that the intensity of the solidarity between theory and experiment varied historically, but he also considered that it could vary synchronically, from

one field of chemistry to another: "Chemistry, epistemologically speaking, did not reach the same point of realistic maturation in each of its parts. One must, therefore, constantly take stock in order to determine how deep the different symbols and schemata are embedded in reality" (Bachelard 1953: 122-123; Eng. tr. is ours). This is why Bachelard claimed that the study of chemistry required a *polyphilosophism* (236-237): a variety of philosophical positions which reflected the plurality of relations between theory and experiment.

However, Bachelard became more sharply critical as he commented on the philosophies that stand at the two extremes of his diagram: idealism and realism. In that text, *idealism* stood for the belief that the laws of the world follow the laws of the mind: the knowledge that the spirit elaborates independently from any relation to experience is by itself knowledge of reality. *Realism* stood for the belief that the mind cannot fully understand reality, that reality irremediably exceeds our cognitive capacities. This definition explains why Bachelard often associated realism with irrationalism, i.e. the depreciation of the powers of reasoning. One reaches these philosophical positions, according to Bachelard, when one completely loses track of either experiment (idealism) or theory (realism); the articulation between these philosophies would consist in an oscillation between two incompatible views, as one shifts from one extreme to another. Bachelard associated such combination of realism and idealism with one philosopher: Émile Meyerson, who can be regarded as Bachelard's designated enemy (Fruteau de Laclos 2015).

Bachelard referred more specifically to *Identity and Reality* (Meyerson 1912; Eng. tr. 2007), where Meyerson claimed that the human mind has one fundamental need: the search for identity, that is, the will to find something which remains permanent through changes. He presented such identification as the most fundamental form of intelligibility and intended to prove that it stood behind most of our intellectual operations such as the structuration of perception and the search for causes of phenomena. Meyerson, however, noted that there is another factor at play in the constitution of knowledge objecting to our quest for an identity: the willingness to report in detail the reality in its changing diversity. Bachelard sure gave a caricatural view of Meyerson's work, as he stated that it pictured a static opposition between the mind, defined by its search for what is identical, and reality, defined by the resistance it opposes to that operation. *Identity and Reality* instead showed how the will to combine these two requirements gave its dynamics to the process of knowledge: the search for identity and what we call reality take different forms, according to Meyerson, which evolve conjointly as new experiments challenge our first attempts to find an identity – which is, one may argue, closer to a Bachelardian dialectics than Bachelard himself would have admitted it.

While Bachelard's reading of Meyerson may be unfair, I propose looking at the extremities of his 1947 diagram from another angle: I will regard the criticism of Meyerson as an indirect form of self-criticism, revealing of Bachelard's own philosophical evolutions. Jean-Claude Pariente (2015) indeed stated that such hesitation between idealism and realism characterized Bachelard's 1927 two doctoral theses (*Thèses de doctorat d'État*), the *Essai sur la connaissance approchée* ("Essay on approximate knowledge") and the *Étude sur l'évolution d'un problème de physique: la propagation thermique dans les solides* ("Study on the evolution of a philosophical problem: thermic propagation in solids"). Even though these two works were written at the same period, there are deep tensions between them, which is probably linked to the fact that they belong to two different genres: the first one belonging to the traditional philosophy of knowledge; the second being a study of the history of physics. In his philosophical *Essay*, Bachelard was deeply influenced by Bergsonism and by Émile Meyerson himself, who can be regarded as part of Bergsonism in a wide sense (Fruteau de Laclos 2009). He made a positive reference to the work of Meyerson and stood much closer to what he later came to designate pejoratively as "naïve realism": "Mr. Meyerson has proved that science commonly postulates a reality. In our point of view, this reality presents, in its inexhaustible unknown, a character which is eminently favorable to the development of an endless research. All its being lays in its resistance to knowledge. We will thus take as the postulate of epistemology the fundamental incompleteness of knowledge." (Bachelard 1928: 13; Eng. tr. is ours). The philosophical essay thus stressed the limitations of scientific knowledge. It regarded mathematical physics as an example of the mind's activity of tracking (*repérage*), which artificially structures reality by neglecting some of its aspects and claimed that it should ultimately give way to a more direct apprehension of reality. The study of the history of mathematical physics, however, has given more power to mathematics: while Bachelard was studying the process that led to the formulation of the Fourier differential equations on heat propagation, he realised that mathematics does not simplify reality, but rather complicates our thinking and is the key to achieving objective knowledge. The first chapters of his study insisted on the fact that the common knowledge of heat was superficial and did not open any path for further progress; only the mathematical search for the equation of propagation would ultimately allow us to understand what heat is and how it works. Bachelard's conclusion was tempted by the adoption of a certain form of idealism – as we found in the introduction to *Le Rationalisme appliqué*, i.e. the assertion that reality follows the rules of the spirit. He insisted on the "prophetic sense of mathematical physics" (Bachelard 1973: 168; Eng. tr. is ours), and claimed that such predictive power

was all the more surprising that, according to him, these mathematical relations follow properly mathematical rules that "are at no time inspired by the relations of reality" (*Ibid.*: 169). He considered that such power of mathematical reasoning could not be accounted for by assuming that mathematical theories merely sum up our empirical knowledge: "We can thus understand to a certain extent that mathematics offers an appropriate language in order to handle general elements that the scientific analysis of phenomena has set apart. However, our astonishment entirely subsists when we see mathematical physics, which Cauchy calls *sublime physics*, get ahead of observation and predict laws which refine and extend experimental laws." (*Ibid.*: 168). The study of the history of mathematical physics, therefore, led Bachelard to insist on what Eugene Wigner (1990) would later call "the unreasonable effectiveness of mathematics in the natural sciences".

Bachelard's philosophical evolutions in the early 1930s can be considered as a way to solve the tensions that existed between his two doctoral dissertations. Since *The New Scientific Spirit* (Bachelard 2020; Eng. tr. 1984), his presentation of a dialectics between theory and experiment can, indeed, be regarded as a way of combining two requirements which seemed incompatible. Bachelard granted that mathematical reasoning does have the power to "get ahead of observation and predict laws which refine and extend experimental laws" (Bachelard 1973: 168), but he denied that it could do so independently from a perpetual dialogue with experiment. On the other hand, he granted that reality could exceed our rational constructions, but denied that this should be regarded as a fundamental limitation of scientific knowledge and rather claimed that this limit manifests itself in the form of concrete technical difficulties or experimental objections, which can only lead to a refinement of our theoretical knowledge. In his 1947 diagram, Bachelard presented realism and idealism as two extreme poles that can be reached when one loses the balance between theory and experiment. It seems that such a presentation inversed the real biographical order, the evolution through which Bachelard could finally overcome his hesitation between idealism and realism and combine them in the form of an applied rationalism and technical materialism.

The introduction to *Le Rationalisme appliqué* can be regarded as one of the most canonical expositions of Bachelard's philosophy of science. It echoes the introductions and conclusions of each of his epistemological works since *The New Scientific Spirit*, in a series of small philosophical treatises which offer a remarkable continuity. The central chapters of his epistemological works put these philosophical conceptions into play, offering an analysis of the latest scientific developments of his time. Bachelard's passionate readings of scientific publications enabled him to investigate the concrete and specific forms of this

dialectics between theory and experiment. His legacy can be found in further attempts to investigate the solidarity between theory and experiment, as illustrated, for instance, the new trend of historical epistemology which developed at the Max Planck Institute for History of Science. Hans-Jörg Rheinberger (2005) praised how the notion of phenomenotechnique "conceptualizes the relation between scientific thinking and technology in modern science" (313), revealing that each scientific object "derives its existence from a circuit that is at the same time material and discursive" (316); he tried to investigate other kinds of material and discursive circuits in his personal work (Rheinberger 1997, 2010). We could also say, using the concepts developed by Lorraine Daston (2000), that the noumenon and the techno-phenomenon are the two sides of a scientific object: its theoretical and experimental mode of existence. We can finally question how these authors relate to Bachelard's work by wondering where they would stand in Bachelard's philosophical diagram: take, for instance, the social constructivism of Latour and Woolgar (1986) or the entity realism of Hacking (1983). Would they stand at the center of Bachelard's diagram, as a new form of his applied rationalism and technical materialism, or should we consider that they bend towards idealism and realism respectively (Vagelli, 2018)? Wondering what legitimate grounds for historians and philosophers could be to acknowledge a preference for one of the sides of Bachelard's diagram seems like the most efficient way to challenge his applied rationalism.

Lucie Fabry
lucie.fabry@ens.psl.eu
École Normale Supérieure, République des Savoirs, PSL University, Paris
and Centre Marc Bloch, Deutsch-Französisches Forschungszentrum
für Sozialwissenschaften, Berlin.

References

Althusser, Louis, 1965, *Pour Marx*, F. Maspero, Paris; Eng. tr. by Ben Brewster 2010, *For Marx*, Verso, London.

Alunni, Charles, 2015, "Gaston Bachelard face aux mathématiques", in *Revue de Synthèse*, 136: 9-32.

Ayer, A.J., 1958, "Philosophie et langage ordinaire", in *Dialectica*, 12, 2: 99-129.

Bachelard, Gaston, 1928, *Essai sur la connaissance approchée*, J. Vrin, Paris.

Bachelard, Gaston, 1932, *L'Intuition de l'instant: étude sur la Siloë de Gaston Roupnel,* Stock, Paris; Eng. tr. by Eileen Rizo-Patron 2013, *Intuition of the instant*, Northwestern University Press, Evanston, Illinois.

Bachelard, Gaston, 1938a, *La Formation de l'esprit scientifique: contribution à une psychanalyse de la connaissance objective*, Librairie philosophique J. Vrin, Paris; Eng. tr. by Mary McAlister Jones 2002a, *The formation of the scientific mind: a contribution to a psychoanalysis of objective knowledge*, Clinamen, Manchester.

Bachelard, Gaston, 1938a, *La Psychanalyse du feu*, Gallimard, Paris; Eng. tr. by Alan C. M. Ross 1968b, *The psychoanalysis of fire*, Beacon Press, Boston.

Bachelard, Gaston, 1939, *Lautréamont*, Librairie José Corti, Paris; Eng. tr. by Joanne H. Stroud 1986, *Lautréamont*, Dallas Institute Publications, Dallas.

Bachelard, Gaston, 1940, *La philosophie du non: essai d'une philosophie du nouvel esprit scientifique*, PUF, Paris; Eng. tr. by G. C. Waterston 1968a, *The Philosophy of No: A philosophy of the new scientific mind*, Orion Press, New-York.

Bachelard, Gaston, 1942, *L'Eau et les rêves: essai sur l'imagination de la matière*, Librairie José Corti, Paris; Eng. tr. by Edith R. Farrell 1983, *Water and dreams: an essay on the imagination of matter*. Pegasus Foundation, Dallas.

Bachelard, Gaston, 1943, *L'Air et les songes: essai sur l'imagination du mouvement*, Librairie José Corti, Paris; Eng. tr. by Edith R. Farrell, Frederick C. Farrell, 1988, *Air and dreams: an essay on the imagination of movement*, Dallas Institute Publications, Dallas.

Bachelard, Gaston, 1945, "La philosophie scientifique de Léon Brunschvicg", in *Revue de Métaphysique et de Morale*, PUF, 50, 1/2: 77-84.

Bachelard, Gaston, 1947, "La philosophie dialoguée", in *Dialectica*, 1, 1: 11-20.

Bachelard, Gaston, 1948a, *La Terre et les rêveries de la volonté*, Librairie José Corti, Paris; Eng. tr. by Kenneth Haltman 2002b, *Earth and reveries of will: an essay on the imagination of matter*, Dallas Institute Publications, Dallas.

Bachelard, Gaston, 1948b, *La Terre et les rêveries du repos*, Librairie José Corti, Paris; Eng. tr. by Mary McAlister Jones 2011, *Earth and reveries of repose: an essay on images of interiority*, Dallas Institute Publications, Dallas.

Bachelard, Gaston, 1949, *Le rationalisme appliqué*, PUF, Paris.

Bachelard, Gaston, 1953, *Le matérialisme rationnel*, PUF, Paris.

Bachelard, Gaston, 1973, *Étude sur l'évolution d'un problème de physique: la propagation thermique dans les solides*, J. Vrin, Paris.

Bachelard, Gaston, 2020, *Le nouvel esprit scientifique*, PUF, Paris; Eng. trad. by Arthur Goldhammer 1984, *The new scientific spirit*, Beacon Press, Boston.

Bachelard, Gaston, Paul Bernays, Ferdinand Gonseth, 1947, "Editorial", in *Dialectica*, 1, 1: 4-8.

Bensaude-Vincent, Bernadette, 2005, "Chemistry in the French tradition of philosophy of science: Duhem, Meyerson, Metzger and Bachelard", in *Studies in History and Philosophy of Science Part A*, 36, 4: 627-649.

Bensaude-Vincent, Bernadette, 2012, "Gaston Bachelard (1884-1962)", *Philosophy of Chemistry*. North-Holland, Amsterdam: 141-150.

Bertholet, Edmond, 1968, *La Philosophie des sciences de Ferdinand Gonseth*, Ed. l'Age d'homme, Lausanne.

Bohr, Niels, 1948, "On the notions of causality and complementarity", in *Dialectica*, 2, 3-4: 312-319.

Bontems, Vincent, 2010, *Bachelard*, Les Belles Lettres, Paris.

Bontems, Vincent, 2013, "L'éthique de l'ouverture chez Gaston Bachelard et Ferdinand Gonseth", *Gaston Bachelard: science et poétique, une nouvelle éthique?*, Hermann, Paris: 379-397.

Bontems, Vincent, 2018, "Le 'non-cartésianisme'. La méthode non-cartésienne selon Gaston Bachelard et Ferdinand Gonseth", in Kolesnik-Antoine, Delphine (ed.), *Qu'est-ce qu'être cartésien?*, ENS Éditions, Lyon: 567-580.

Bourdieu, Pierre, Jean-Claude Chamboredon, Jean-Claude Passeron, 1972, *Le Métier de sociologue* : préalables épistémologiques, 2nd edition, Mouton, Paris, La Haye; Eng. tr. by Richard Nice 1991, *The Craft of Sociology: Epistemological Preliminaries*, Walter De Gruyter, Berlin, New York.

Broglie, Louis de, 1948, "Sur la complémentarité des idées d'individu et de système", in *Dialectica*, 2, 3-4: 325-330.

Brunschvicg, Léon, 1922, *L'expérience humaine et la causalité physique*, F. Alcan, Paris.

Canguilhem, Georges, 1965, *La Connaissance de la vie*, J. Vrin, Paris; Eng. tr. by Paola Marrati-Guénoun 2008, *Knowledge of life*, Fordham University Press, New York.

Carnap, Von Rudolf, 1958, "Beobachtungssprache und theoretische Sprache", in *Dialectica*, 12, 3-4: 236-248.

Castelão-Lawless, Teresa, 1995, "Phenomenotechnique in historical perspective: its origins and implications for philosophy of science", in *Philosophy of Science*, 62, 1: 44-59.

Chimisso, Cristina, 2008, "From Phenomenology to Phenomenotechnique: The Role of Early Twentieth-Century Physics in Gaston Bachelard's Philosophy", in *Studies in History and Philosophy of Science Part A*, 39, 3: 384-392.

Dagognet, François, 1965, "Brunschvicg et Bachelard", in *Revue de Métaphysique et de Morale*, 70, 1: 43-54.

Donatiello, Paola, Francesco Galofaro, Gerardo Ienna, 2018, *Il senso della tecnica: saggi su Bachelard*, Esculapio, Bologna.

Duhem, Pierre Maurice Marie, 2006, *La Théorie physique. Son objet, sa structure*, ENS Éditions, Lyon; Eng. tr. by Philip P. Wiener 1991, *The Aim and Structure of Physical Theory*, Princeton University Press.

Dupréel, Eugène, 1957, "Consistance et valeurs", in *Dialectica*, 11, 3-4: 345-353.

Einstein, Albert, 1948, "Quanten-Mechanik und Wirklichkeit", in *Dialectica*, 2, 3-4: 320-324.

Fabry, Lucie, 2019, "Phenomenotechnique: Bachelard's critical inheritance of conventionalism", in *Studies in History and Philosophy of Science Part A*, 75: 34-42.

Fruteau de Laclos, Frédéric, 2009, *Le cheminement de la pensée selon Émile Meyerson*, Presses universitaires de France, Paris.

Fruteau de Laclos, Frédéric, 2015, "Le bergsonisme, point aveugle de la critique bachelardienne du continuisme d'Émile Meyerson", in Wunenburger, Jean-Jacques, Frédéric Worms (eds), *Bachelard et Bergson : continuité et discontinuité*. Presses Universitaires de France, Paris: 109-122.

Gaukroger, Stephen W., 1976, "Bachelard and the Problem of Epistemological Analysis", in *Studies in History and Philosophy of Science Part A*, 7, 3: 189-244.

Gödel, Kurt von, 1958, "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes", in *Dialectica*, 12, 3-4: 280-287.

Gonseth, Ferdinand, 1936, *Les Mathématiques et la réalité, essai sur la méthode axiomatique*, Félix Alcan, Paris.

Gonseth, Ferdinand *et al.*, 1939, "Présentation et défense de l'idonéisme", in *Revue de Théologie et de Philosophie*, Librairie Droz, 27, 113: 260-287.

Goodstein, Reuben Louis, 1958, "On the nature of mathematical systems," in *Dialectica*, 12, 3-4: 296-316.

Granger, Gilles Gaston, 1983, *Formal thought and the sciences of man*, D. Reidel Publishing Company, Dordrecht, Holland.

Granger, Gilles-Gaston, 1987, "Le rationnel selon Gaston Bachelard", in Lafrance, Guy (ed.), *Gaston Bachelard: profils épistémologiques*, Presses de l'Université d'Ottawa, Ottawa: 9-23.

Hacking, Ian, 1983, *Representing and intervening: introductory topics in the philosophy of natural science*, Cambridge University Press, Cambridge.

Heisenberg, Werner, 1948, "Der Begriff abgeschlossene Theorie in der modernen Naturwissenschaft", in *Dialectica*, 2, 3-4: 331-336.

Jung, Carl Gustav, 1951, "Grundfragen der Psychotherapie", in *Dialectica*, 5, 1: 8-24.

Lamy, Julien, 2005, "Enquête sur le concept de 'noumène' dans l'épistémologie bachelardienne. La physique contemporaine comme science nouménale", in *Cahiers Gaston Bachelard*, Centre Georges Chevrier (CGC), Université de Bourgogne, 7: ???

Latour, Bruno, Steve Woolgar, 1986, *Laboratory life: the construction of scientific facts*, Princeton University Press, Princeton, N.J.

Meyerson, Emile, 1912, *Identité et réalité*, 2nd edition, Felix Alcan, Paris; Eng. tr. by Kate Loewenberg 2007, *Identity & reality*, Routledge, Oxfordshire; New York.

Pariente, Jean-Claude, 2015, "Rationalisme et ontologie chez Gaston Bachelard", *L'épistémologie française, 1830-1970*, Éditions matériologiques, Paris: 235-263.

Pauli, Wolfgang, 1948, "Editorial", in *Dialectica*, 2, 3-4: 307-311.

Piaget, Jean, 1950, "Epistémologie génétique et méthodologie dialectique II", in *Dialectica*, 4, 4: 287-295.

Piaget, Jean, 1954, "Inconditionnés transcendantaux et épistémologie génétique", in *Dialectica*, 8, 1: 5-13.

Piaget, Jean, 1959, "Perception, apprentissage et empirisme", in *Dialectica*, 13, 1: 5-15.

Popper, Karl, 1978, "Natural selection and the emergence of mind", in *Dialectica*, 32, 3-4: 339-355.

Ramond, Charles (ed.), 2011, *Spinoza: nature, naturalisme, naturation*, Presses universitaires de Bordeaux, Pessac.

Reichenbach, Hans, 1948, "The Principle of anomaly in Quantum Mechanics", in *Dialectica*, 2, 3-4: 337-350.

Rheinberger, Hans-Jörg, 1997, *Toward a history of epistemic things: synthesizing proteins in the test tube*, Stanford University Press, Stanford.

Rheinberger, Hans-Jörg, 2005, "Gaston Bachelard and the Notion of 'Phenomenotechnique'", in *Perspectives on Science*, 13, 3: 313-328.

Rheinberger, Hans-Jörg, 2010, *An epistemology of the concrete: twentieth-century histories of life*, Duke University Press, Durham.

Robinet, André, 1961, "La philosophie malebranchiste des mathématiques", in *Revue d'histoire des sciences*, 14, 3: 205-254.

Ruyer, Raymond, 1959, "La Psychobiologie et la science", in *Dialectica*, 13, 2: 103-122.

Skolem, Thoralf, 1958, "Reduction of axiom systems with axiom schemes to systems with only simple axioms", in *Dialectica*, 12, 3-4: 443-450.

Spinoza, Baruch, 2020, *Ethica*, PUF, Paris.

Vagelli, Matteo, 2018, "Bachelard, Hacking e il realismo tecnoscientifico", in *Il senso della tecnica: saggi su Bachelard*, Esculapio, Bologna: 121-136.

Vinti, Carlo, 1997, *Il soggetto qualunque: Gaston Bachelard fenomenologo della soggettività epistemica*, Ed. scientifiche italiane, Napoli.

Wigner, Eugene P., 1990, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", in Mickens, Ronald E. (ed.), *Mathematics and Science*, World Scientific, Singapore: 291-306.

The dialogical philosophy. La philosophie dialoguée[1]

Gaston Bachelard

Abstract: This essay points out the necessity of evoking several philosophical systems in order to realize the evolution of the scientific theory of knowledge in modern physics. It proposes a sort of spectrum of philosophical systems with seven conceptions set in the following order : realism, empiricism, positivism, rationalism, formalism, conventionalism, idealism. A double filiation unites these philosophies in the center of the spectrum, so that, rationalism, in conjunction with technical materialism, seems to be the most strongly established philosophy, and the backbone of modern scientific thought. Rationalism, far from representing a detached point of view, appears as a dialectical philosophy as soon as it seeks its confirmation in technical experience.

I

If we follow with attention, that is, with passionate interest, the activity of contemporary physics, we see the development of a philosophical dialogue which has the merit of being exceptionally precise: the dialogue between the experimenter provided with precise instruments and the mathematician who aspires to closely inform the experiment. Whereas, too often, in philosophical debates realists and rationalists do not manage to talk of *one same thing*, we have the neat and comforting impression that, in the scientific dialogue, both the interlocutors speak of *the same matter*. Whereas in philosophy conferences, we see philosophers exchanging *arguments*, in the conferences of physics, we see experimenters and theoreticians exchange *information*. Is it not necessary that the experimenter be informed about the theoretical aspect of data that the mathematician judges to be highly coordinated, without which the experimenter's interpretations can fall victim to his own personal views? And is it not necessary as well that the theoretician be informed on

[1] Originally published in *Dialectica*, vol. 1, issue 1, "What is Dialectic / L'idée de dialectique / Die dialektische Denkweise" (15/2/1947): 11-20, <http://www.jstor.org/stable/42963796>, published by Wiley and released under CC4.0 NC-BY license.

philinq IX, 1-2021, pp. 231-240
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v9i1.363

all the circumstances related to the experiment, without which his syntheses could remain partial and purely abstract? Physics has then two philosophical poles. It is a true *field of thought* which is specified in mathematics and experiments and comes most to life in the convergence of mathematics and experience. As a strong synthesis, physics determines an *abstract-concrete* mentality. Throughout this work we will try unceasingly to characterize this mentality according to its double action of abstraction and concretisation, by never breaking the connecting mark imposed by language – in the absence of a knowledge of more unitary principles – in order that we can *understand the reciprocity of the dialectics* which move along an endless and two-way path from the mind towards things.

The contact between *experience* and *mathematics* develops as a propagating solidarity. When it is the experimenter who brings the first message of a new phenomenon, the theoretician does not rest until he has modified the prevailing theory in order that it can assimilate the new fact. Through this – undoubtedly late – modification the mathematician shows that theory, now softened, *should have envisaged* the innovation. He likes to make a display of a sort of *recurrent fecundity* which is – as we will show – an important feature of rationalism, since this recurrent fecundity constitutes the foundation of *rational memory*. This memory proper to reason, this memory of coordinated ideas, obeys psychological laws completely different from those of the *empirical memory*. These ideas, put in order, reordered and coordinated within the logical time, determine a veritable emergence of memory. Certainly, nobody – and the experimenter even less so – laughs at this return, afterwards, to the sources of the theoretical prevision. On the contrary, the experimenter is pleased that his discovery is assimilated by mathematics. He knows that a new fact, when connected to the modern aspect of the prevailing theory, is guaranteed by a objectivity that is thoroughly overseen, given that the prevailing theory is a system of experimental examination which is active in the brightest brains of the epoch. We have the impression that the phenomenon is properly seen insofar as it *could have been foreseen*. The theoretical perspective *places* the fact where it is supposed to be. If the fact is correctly assimilated by the theory, there is no more hesitation about the place that it *should* occupy in a *thought*. It is no longer a heteroclite fact, a raw fact. It becomes now a *cultural fact*. It has a rationalist status. It is henceforth the subject of a dialogue between the rationalist and the empiricist.

When it is the theoretician who announces the *possibility* of a new phenomenon, the experimenter addresses this perspective, provided that he feels this latter is aligned with modern science. This is why, at the beginning of the wave mechanics of the electron, one searched for a phenomenon

that, in the case of the electron, could correspond to the phenomenon of light polarization. Whenever such a specific investigation ends in vain, it has nevertheless a positive character for epistemology, since it contributes to the limitation and definition of analogies. Experience thus associated with theoretical views has nothing in common with occasional research, with these experiments "to see" which have no place within strongly structured sciences such as physics and chemistry, within sciences too for which the *instrument* is the intermediary necessary for examining a truly instrumented phenomenon, designated as the object of a phenomenotechnique. No physicist would spend "his credit" to build an instrument with no theoretical destination. In physics, Claude Bernard's experiment "to see" is meaningless.

What tacit agreement reigns in the *city of physics*! In what manner the unrepentant dreamers wanting to "theorize" far from mathematical methods are dismissed! The theoretician must actually possess all the *mathematical past* of physics, that is to say, all the rationalist tradition of experience. The experimenter, on his side, must know entirely the *present of technique*. We would be surprised if a physicist used the old vacuum air pump, even if it was provided with the Babinet tap. Modernism of the technical reality and rationalist tradition of every mathematical theory: this is the double cultural ideal that should permeate all the themes of scientific thought.

The philosophical cooperation of these two aspects of physical science – the rational aspect and the technical aspect – can be synthesized in the following double question:

Under what conditions is it possible to give a *reason* for a *precise* phenomenon? Moreover, the word *precise* is essential, for precision is the sphere of *reason*'s engagement.

Under what conditions is it possible to provide *real* evidence of validity for a mathematical organisation of physical experience?

The time is long past since epistemology considered mathematics as a mere instrument to express the laws of physics. Mathematics of physics are "more committed". It is not possible to *found* physical sciences without entering into the philosophical dialogue between the rationalist and the experimenter, nor is it possible without answering the two – somehow *reciprocal* – questions that we have just set. In other words, the modern physicist needs a double certainty:

1. the certainty that reality is directly related to rationality, so that it can obtain the name of *scientific reality*;
2. the certainty that the rational arguments which concerns experience are already moments of this experience.

To put it simply, no rationality without target, no disjointed empiricism (*pas de rationalité à vide, pas d'empirisme décousu*): these are the two philosophical

obligations which found the strict and precise synthesis of theory and experience in contemporary physics.

This *bi-certainty* is essential. If one of the two terms is missing, we can still do experiments and we can still do mathematics, but we cannot participate in the scientific activity of contemporary physical science. This bi-certainty cannot be expressed but through a two-way philosophy, through a dialogue. Nevertheless, this dialogue is so tight that we cannot detect any character of the old philosophical dualism. It is no longer about bringing a solitary mind face to face with the indifferent universe. From now on one must place oneself in the middle, where the knowing mind is determined by the precise object of its knowledge and where, in return, it determines its own experience with greater precision. It is precisely this *central* position which allows the dialectics of reason and technique to reach its effectiveness. We will try to place ourselves in this central position where an *applied rationalism* and an *instructed materialism* arise as well. Thereafter we will also insist on the power of application proper to every scientific rationalism, that is, to every rationalism that can bring its evidence of fecundity up into the organisation of technical thought. Precisely through this application rationalism achieves its objective values. In this sense, the evaluation of scientific thought no longer lies in a formal, abstract, universal rationalism. It is necessary to achieve a concrete rationalism, in solidarity with increasingly particular and precise experiments. This rationalism must also be sufficiently open to receive new determinations from experience. In experiencing this dialectics a little more closely, we are convinced of *the eminent reality of the fields of thought*. It is within these epistemological fields that the exchange of values between rationalism and experimentalism takes place.

II

In fact, the criss-crossing of two opposite philosophies active within scientific thought involves even more philosophies and we should present dialogues which are undoubtedly less tight but that extend the psychology of the scientific mind. For instance, we would mutilate the philosophy of science, if we did not examine how *positivism* and *formalism* are situated, given that undoubtedly they both have proper functions in contemporary physics and chemistry. Nevertheless, one of the reasons why we believe in the validity of our central position is that all philosophies of scientific knowledge are ordered starting from *applied rationalism*. It is barely necessary to comment on the following table, when we apply it to the scientific thought:

Idealism
↑
Conventionalism
↑
Formalism
↑
Applied rationalism and technical materialism
↓
Positivism
↓
Empiricism
↓
Realism

We only indicate the two perspectives of *weakened* thoughts which, on one hand, lead from rationalism to naive idealism and, on the other hand, from the technical materialism to naive realism.

Therefore, when rational knowledge is systematically interpreted as the constitution of certain *forms*, as a mere equipment of *formulae* suitable to *inform* any kind of experience, then a *formalism* is established. If at all, this formalism can receive the *outcomes* of rational thought, but it cannot do all the work of the rational thought. Moreover, we do not always limit ourselves to formalism. We started a philosophy of knowledge which weakens the role of experience. We are very close to considering theoretical science as a set of *conventions*, a series of more or less *convenient* thoughts organized according to the clear language of mathematics, which however become no more than an *Esperanto* of reason. The convenience of conventions does not remove their arbitrariness from them. These formulae, these conventions, this arbitrariness, we will come quite naturally to submit them to an activity of the thinking subject. So, we approach an idealism. This idealism is no longer admitted in contemporary epistemology but it played such a great role in the philosophies of nature in the 19th century that it must be taken into account in a general examination of the philosophical approaches to science.

Besides, we have to underline the powerlessness of idealism to reconstitute a modern version of rationalism, an active rationalism able to inform the knowledge resulting from the new areas of experience. In other words, we cannot invert the perspective that we have just outlined. In fact, when the idealist establishes a philosophy of nature, he limits himself to ordering the images that he creates of nature, indulging in the immediate aspect of these images. He does not go beyond the limits of an ethereal sensualism. He does not un-

dertake a thorough experience. He would be astonished if asked to follow the inquiries of science into essentially instrumentalist experimentation. He does not think he should have to accept the *conventions* of other minds. He would not consent to that slow discipline which intends to *form* his spirit on the basis of the lessons of objective experience. Idealism misses every opportunity to account for modern scientific thought. Scientific thought cannot reach its sound and multiple forms in such a solitary environment, in this solipsism which represents the congenital sickness of idealism. Scientific thought needs a social reality, the agreement of a city of physics and mathematics. We should then rather place ourselves at the central position, that of *applied rationalism*, working on the institution of a specific philosophy proper to scientific thought.

Viewing our table from the other perspective, instead of the evanescence leading to idealism, we find a progressive inertia of thought, which leads to realism, to a conception of reality as synonym of irrationality.

In fact, when we pass from rationalism, in which physical experience is in strong solidarity with theory, to *positivism*, we have the impression of suddenly losing all the principles of *necessity*. Thereafter, pure positivism is no longer able to justify the power of deduction which is active in the development of modern theories; it cannot account for any of the *values of coherence* proper to contemporary physics. This notwithstanding, with respect to pure empirism, positivism appears to be at least the guardian of the hierarchy of laws. It maintains the right to discard sharp approximations, details and varieties. However, this hierarchy of laws does not have the same value as the organisation of necessities clearly understood by rationalism. Moreover, since it is based on judgements of utility, positivism already tends to *pragmatism*, to that hodgepodge of recipes represented by *empiricism*. Positivism is not at all provided with what is necessary to determine the orders of approximation, to feel that strange sensitivity of rationality given by second-order approximations, that is, this more inexact, controversial and consistent knowledge which we achieve through the accurate examination of minute experiments and which helps us understand that there is greater rationality in complexity than in simplicity.

Moreover, going a step further than empiricism, which loses itself in the story of its own achievements, we reach that pile of facts and things with which *realism* is stuffed, and which gives this latter the illusion of richness. We will show later how far from any scientific mind is the postulate, very easily accepted by some philosophers, which considers reality as a pole of irrationality. When we have led the philosophic activity of scientific thought back to its active center, it will be clear that the function of active materialism is precisely to limit what can be qualified as irrational within its arguments and objects.

Chemistry, fortified by its rational *a-priori*, delivers *substances devoid of accidents*, removing from any material the irrationality of its origins.

We will, however, take this discussion up again on the basis of particular examples. We actually think that some precise examples borrowed from scientific knowledge can make general philosophical discussions *more aware* so long as we do not tackle discussions starting from fixed philosophical convictions. What we intended to present through this quick *philosophical topology* is the frame within which most philosophical discussions about science take place. One feature captures our attention: the different philosophical tones that we have mentioned together form a veritable "spectrum" (*spectre*). In this sense, we intend to say that they quite naturally take on a linear order. In light of new philosophical nuances, it will suffice to open this spectrum up a little more and without needing to modify the order of the fundamental philosophies. On the other hand, if we undertook a similar investigation into the elements of a polyphilosophy for other sciences, such as mathematics, biology, sociology and psychology, then we should certainly determine other spectra for philosophical analysis. Nevertheless, no spectrum is more extensive than the one which helps us class the philosophemes of the physical sciences. Undoubtedly, not all the parts of a science are at the same level of philosophical maturity. It is therefore always concerning precise experiences and problems that the philosophical values of a science have to be determined.

III

If we attempt to philosophically characterize the active scientific notions, we will see that each has two sides, always two sides. Every precise notion is a notion that has been given the character of precision. It was precisely clarified through an effort of 'idoneism', according to the meaning given to this term by Gonseth: this idoneism has only become more advanced as the dialectics have become tighter. However, these dialectics already arise from the extreme symmetries of the table that we have proposed. In this sense, we could already clarify the problems facing the epistemology of the physical sciences y of science, if we established the 'dialogical philosophy' of formalism and positivism, epistemological doctrines which are equidistant from the center of the most strongly coordinated thoughts. Formalism would then coordinate with sufficient clarity all the mathematical perspectives which inform the positive laws provided by scientific experience. Without having the apodicticity of rationalism, formalism is provided with logical autonomy.

It would still be possible to detect connections between empiricism and conventionalism: philosophies which are both undoubtedly too loose. Their

dialogue would at least have the charm of a double-skepticism. They are thus generally appreciated by modern philosophers, who observe from afar the progress of scientific thought.

As for the two extreme philosophies, idealism and realism, their dogmatism is their only strength. In particular, it is difficult to explain how scientific realism might emancipate itself from common realism. If science were the description of a given reality, what would ever entitle it to organize this description.

Our task will therefore be to show that rationalism is not at all in solidarity with the imperialism of the subject, that it cannot develop in an isolated consciousness. We shall also demonstrate that technical materialism essentially corresponds to a transformed reality, a rectified reality: a reality which has been granted precisely the ultimate human mark, the mark of rationalism.

So, we will always be brought back to the philosophical center, which is the basis of reflective experience as well as of rational invention, in other words, back to that region in which contemporary science actually operates.

IV

In these conditions, a philosophy like that of Émile Meyerson which, by appealing to two poles *apart from each other*, determines the savant's simultaneous attachment to reality and to the identical, does not seem to give rise to an epistemological field of sufficient intensity. To consider the savant as both an absolute realist and a rigorous logician leads us to juxtapose general philosophies that are ineffective. These are not philosophies at work, but rather *summary*-philosophies which can only contribute to the characterisation of historical periods of time. Through technical progress, the "reality" examined by the savant changes its appearance and loses the character of permanence which is the basis of philosophical realism. For instance, the "electric reality" of the 19th century is much different from the "electric reality" of the 18th.

On the other hand, barely has a reduction to the identical been made when again the research for diversity starts from the identical, so that it will be necessary to unceasingly revive the dialectics of what is identified and what is diversified. Reality, as well will be concerned by a multiplication of the dialectics of analysis and synthesis, of pruning and construction, selection and realisation. A science that is continually rectified in its principles and subjects cannot be granted a unitary philosophical designation. It is dialectics not only because of the detail of its approach, but also for the double ideal of its theoretical coherence and experimental precision.

It is probably no doctrinal accident that led Meyerson to a *static conception* of the psychology of the scientific mind. If one believes that the state of mind

of a pre-Lavoisian chemist, such as Macquer, can be similar to the state of mind of a contemporary chemist, one remains confined to an unmoving materialism, a materialism without dialectics. History of science, in this sense, is often deceptive. It almost never conveys the obscurity of thought. It cannot then grasp the rationality as it takes shape. Our current knowledge clarifies in such a vivid way the past of scientific thoughts that we may take every glimmer for actual lights. One thus believes in a reason constituted without an effort of rationality. Léon Brunschvicg saw the weakness of such an absolutist position was and often insisted on the essential relativity of reason and experience: "We lose touch with the course of reality… with that knowledge, whenever we insist on pushing *rationality* and *objectivity* outside ourselves, and end up isolating and opposing the double entity of *absolute reason* and *absolute object*." As we will see, it is by systemically developing a dialectics of cooperation between reason and scientific object that we will best obtain the rational characteristics of technical materialism and, vice versa, the real characteristics of applied rationalism. Here again, what provides guarantees concerning the object are not the primary experiences but the sharp approximations. Considered in relation to its applications, a rational organisation of experience is not merely the *aim* (*visée*) of a mind which would be enlightened by the mere awareness of the identity of his apperceptions. The intentionality of applied rationalism holds the possibility of self-rectification in reserve. In its application, it is open to those dialectics that can produce resonances up to the principles of organisation. In other words, the second approximation has not the same epistemological structure as the first. It is at the level of the second approximation that dialectics are truly active. These dialectics associate the mathematical mind (*esprit de géométrie*) with the intuitive mind (*esprit de finesse*) into a synthesis which is clearly active in the contemporary scientific mind.

Epistemology must then be as dynamic as science. By multiplying the number of reciprocal figures that we called *Brunschvicg's doublets*[2], we hope to bring together the coherence of rational thought and the *cohesion* of technical materialism. This notwithstanding, the several doublets composed by Brunschvicg according to the Spinozian pattern of *natura naturans* and *natura naturata*, such as *spatializing space* and *spatialized space*, *numbering number* and *numbered number* have to become even *more tightly bound*, in order to account properly for the strong *coupling* of ideas and experiences that arises from the development of contemporary physics and chemistry.

[2] Bachelard, Gaston, 1945, "La philosophie scientifique de Léon Brunschvicg", *Revue de Métaphysique et de Morale*, 50 (1/2): 77-84 <http://www.jstor.org/stable/40899137>, quot. p. 81.

The epistemologist will have to apply the dialogical philosophy to doublets borrowed above all from physics and chemistry, since these doublets allow the traditional debate on the realness of the sensible world to become more precise. However, there will be many occasions to slightly shift the debate. This will be the case, for instance, in the debate on the duality of *symbolising symbol* and *symbolized symbol* in organic chemistry. There is, in fact, a remarkable epistemological difference between the *symbols* that aim only at intuitively translating a general knowledge and the *models* within which a more realist and more particular knowledge emerges. The conventionalism of the early representations, as they were proposed in the 19th century, has been replaced by a technical materialism which *realizes* schemata.

Likewise the objectifying tendency of rational mind is so strong that, in the mathematics aiming at the proliferation of the abstract, it is not impossible to detect structures which may refer to an objective study. There is therefore room for a post-abstractive experience. Of course, we must regard the empiricism which likes to place procedures for surveying land at the basis of geometry as liquidated. Such references serve no purpose in a modern culture, they can be even dangerous, if their naivety is not corrected as soon as possible. In fact, the subject must be constituted according to rationality, and reach *principles of necessity*. In geometry, demonstration is not about showing but proving. It is this kind of *emergence* that precisely occurs in contemporary physical sciences. Values completely different from convention and observation emerge in natural sciences. The philosopher who intends to follow the life of scientific thought in detail, will come to know the extraordinary couplings of Necessity and Dialectics.

Gaston Bachelard.

Translated from the French by Gennaro Lauro[3]

[3] We would like to thank Samuel Talcott for his accurate reading and invaluable advice.