philosophical inquiries
www.philinq.it

philinq

philosophical inquiries

XII, 1
2024

Edizioni ETS

Essays published on “Philosophical Inquiries” are double-blind peer-reviewed.

Please visit www.philinq.it for submissions and guidelines.

six-monthly journal / periodico semestrale
Subscription (paper, individual): Italy € 50,00, Abroad € 80,00
Subscription (paper, institution): Italy € 60,00, Abroad € 100,00

Subscription fee payable via Bank transfer to
Edizioni ETS
Intesa San Paolo
IBAN IT 21 U 03069 14010 100000001781
BIC/SWIFT BCITITMM
reason: abbonamento “Philosophical Inquiries”

Registrazione presso il Tribunale di Pisa n. 1/13

Direttore responsabile: Alessandra Borghini

EDIZIONI ETS
Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com
www.edizioniets.com
Finito di stampare nel mese di febbraio 2023

Distribuzione
Messaggerie Libri SPA
Sede legale: via G. Verdi 8 - 20090 Assago (MI)

Promozione
PDE PROMOZIONE SRL
via Zago 2/2 - 40128 Bologna

ISBN 978-884677108-7
ISSN 2281-8618

Table of Contents

Essays

Brent Delaney
Hans Vaihinger's Hume: Real fictions and semi-fictions in the Treatise 9

Filippo Nobili
Consciousness lived through. Husserl's reception of Brentano's inner consciousness 37

Simone Cuconato
Can nonexistent mathematical objects make a difference? Meinongianism, indispensability argument and mathematical entanglement 61

Focus

Pierluigi Barrotta, Gustavo Cevolani, Roberto Gronda, Luca Tambolo
Introduction: The role of experts in democratic societies: in honor of Paul K. Feyerabend 81

John Preston
Feyerabend, experts, and dilettantes 89

Pierluigi Barrotta, Roberto Gronda
Wisdom, scientific expertise, and laypeople: some remarks on Paul Feyerabend's philosophy of expertise 107

Karim Bschir
Feyerabend's humanitarian pluralism and its relevance for science-based policy 125

philinq XII, 1-2024
ISSN (print) 2281-8618-ETS

Luca Tambolo, Gustavo Cevolani
Feyerabendian pluralism in practice: lessons from the Di Bella case 145

Piero Avitabile, Alessandro Demichelis
Expanding epistemic public trust.
What role for expert-lay communication? 163

Essays

Hans Vaihinger's Hume: Real fictions and semi-fictions in the *Treatise*

Brent Delaney

Abstract: Hans Vaihinger prefaces the English edition of his seminal text, *The Philosophy of As If*, by drawing attention to his philosophical predecessors. While Hobbes and Berkeley are afforded due credit in the development of philosophical fictionalism, Hume is conspicuously absent. I argue that Hume's early theory of fiction, which he abandoned after publishing the *Treatise*, prefigures Vaihinger's central distinction between two types of fiction: real fictions and semi-fictions. To that end, Hume's philosophy offers a significant contribution to the history of early modern fictionalism.

Keywords: David Hume, Hans Vaihinger, Fiction, Fictionalism, Natural Belief

In 1911, Hans Vaihinger published his *Die Philosophie des Als Ob*, a work translated into English by C.K. Ogden as *The Philosophy of 'As if': A System of the Theoretical, Practical and Religious Fictions of Mankind*. Based on his dissertation completed in 1877, the systematic philosophical account presents a general theory of fiction. Specifically, Vaihinger offers a comprehensive analysis of the instruments of thought in so far as thought is considered a purposive organic function of the psyche. Although Vaihinger's system received less attention after the Second World War, his philosophy of 'Critical Positivism' has become deeply relevant to contemporary research in the last forty years. Most notably, several recent species of philosophical fictionalism may be traced to seminal ideas in Vaihinger's text, namely, scientific, religious, moral, mathematical, and metaphysical fictionalism.

In the Preface to the English edition of *The Philosophy of As If*, Vaihinger surveys a list of philosophers he believes to be his forerunners. First, he draws attention to the movement of English Nominalism, which, he says, represents an initial understanding of fictions. Specifically, he mentions the skeptical philosophical approach of John Duns Scotus, followed by the theory of *ficta* developed by William of Occam. For Vaihinger, Occam offers the first "clear and definite treatment of the fictional nature of general ideas, developed in a

philinq XII, 1-2024, pp. 9-36
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.411
Submitted: 10 April 2022
Accepted: 5 January 2024

manner which is still a model for to-day" (*PAI*: vii).[1] Occam's account introduces the thesis that *ficta* ought to be regarded as practically necessary ideas despite their theoretical non-existence.

Second, while Vaihinger acknowledges George Berkeley to have had a minimal understanding of fictions, it is Hobbes, he argues, who demonstrated "considerable knowledge both of Fictions themselves and of the theory of their use. Empty space, the idea of a *bellum omnium contra omens*, and of an 'original contract' are for Hobbes conscious Fictions" (*PAI*: vii). Vaihinger also names Jeremy Bentham and Adam Smith as noteworthy predecessors.

The main source of inspiration for Vaihinger is, not surprisingly, Immanuel Kant. Vaihinger remains an important figure in the Neo-Kantian tradition – he founded the academic journals, *Kant-Studien* and *Kant Gesellschaft* in 1896 and 1901, respectively. Vaihinger maintains that Kant devoted "100 pages" of his work to outlining a theory of fiction in which he proved "a large number of ideas, not only in metaphysics but also in mathematics, physics and jurisprudence, were Fictions" (*PAI*: viii).

The only additional English philosophers that Vaihinger mentions in his Preface are David Hume and Francis Bacon. He claims that both Bacon and Hume failed to understand fictions as practical necessities of thought. However, Vaihinger concedes that his first major influence was, in fact, Hume: "it was David Hume and still more J.S. Mill whose influence on my thought was paramount" (*PAI*: vii). He then reiterates that his "private studies were devoted mostly to David Hume and John Stuart Mill, whose exact knowledge was decisive for [his] philosophic attitude" (*PAI*: xxxv).

Now, given that Hume specifically refers to the concept of 'fiction' on fifty separate occasions in the *Treatise*, it is odd why Vaihinger does not credit him with even a minimal understanding of pragmatic fictions. Even more curious is why Vaihinger singled out Hobbes for praise and not Hume, considering Hobbes' plausible influence on Hume. Paul Russell has drawn attention to "significant affinities between" Hume's theory of spatial representation and Hobbes' (Russell 2008).[2] Jonathan Cottrell agrees, suggesting that when

[1] '*PAI*' are to *The Philosophy of As If* (Vaihinger 1935), followed by the page number. References to Hume are cited as follows: '*T*' are to *A Treatise of Human Nature*; followed by Book, part, section, paragraph (Norton and Norton 2007), and then corresponding page number in the 1978 Selby-Bigge edition revised by Nidditch: 'SBN'. '*E*' are to *An Enquiry concerning Human Understanding*, followed by section, paragraph (Millican 2007), and then corresponding page number in the 1975 Selby-Bigge edition revised by Nidditch: 'SBN'; '*D*' are to *Dialogues concerning Natural Religion*, followed by section and paragraph (Coleman 2007).

[2] Russell specifically writes: "Hobbes and Spinoza are rarely, if ever, associated with Hume's position on space – which is especially surprising in the case of Hobbes, as there are significant affinities between their views" (2008: 110).

"Hobbes claims that the 'place' (*locus*) of a body is 'feigned' (*ficta*)… he seems to mean that being located involves a relation between a body and our minds… [that] a body has 'feigned' location means that it is related to our minds in a certain way: namely, that it produces a certain effect in our imagination" (2016: 69-70). For Cottrell, Hobbes' treatment of feigned location shares important similarities with Hume's rendering of fictitious duration and distance. Likewise, there seems to be a further point of contact between Hobbes' rendering of the state of nature and Hume's recognition of it as a philosophical *fiction*. In view of these salient connections, why does Vaihinger not acknowledge any tie between the two philosophers, especially when he cites Hobbes' empty space, the war of all against all, and the original contract *as* fictions?

As is well known, Hume was also influenced by George Berkeley, and while it is unclear to what extent Hume was influenced by Berkeley's criticisms of mathematics specifically, it is at least plausible that Hume's mathematical views were minimally informed by Berkeley's controversial critique of the discipline. Vaihinger says of Berkeley that he "proved, quite correctly and with wonderful insight, that practically all the fundamental principles of mathematics were contradictory. From this he drew the conclusion that the mathematicians had no right whatsoever to scoff at the incomprehensible elements and mysteries of Christianity, since their own subject had the same defects" (*PAI*: 117). Hume equally reveals the contradictory nature of important mathematical concepts, namely, equality, unity, infinite divisibility, and identity.[3] As a consequence, Berkeley's *The Analyst*, published only five years before the *Treatise*, may have been a source of influence for Hume's theory of fiction. Vaihinger, though, fails to notice the potential connection between them.

Not only that, in several instances Vaihinger appears to deny Hume a charitable reading. For instance, Vaihinger argues that "when Hume called the categories fictions, he was right *in fact*, though his idea of a fiction was very different from ours. His idea of the 'fiction of thought' was that of a merely subjective fancy, while ours (borrowed from the usage of mathematics and jurisprudence) includes the idea of utility. This is really the kernel of our position, which distinguishes it fundamentally from previous views" (*PAI*: 99). Vaihinger is, of course, incorrect here, for while Hume does believe some fictions are subjective fancies, he believes that other fictions are naturally constitutive of human nature, and

[3] Compare Vaihinger's chapter "The Fiction of the Infinitely Small" (*PAI*: 236-245) with Hume's treatment of infinite divisibility. Hume argues, for instance, that "'tis evident, that as no idea of quantity is infinitely divisible, there cannot be imagin'd a more glaring absurdity…And as this absurdity is very glaring in itself, so there is no argument founded on it, which is not attended with a new absurdity, and involves not an evident contradiction" (*T* 1.2.4.32, SBN 52).

that they are necessarily tied to the idea of utility.[4] Contra Vaihinger, Saul Traiger argues Hume's understanding of fictions may have indeed been influenced by the notion of a legal fiction. Because Hume had studied law, he was "well versed in legal fictions. He explicitly refers to and describes legal fictions in *The History of England*…The notion of legal fiction in Hume's time was, and remains today, a fundamental concept in the law. Its application to Hume's metaphysics and epistemology is natural and fitting, and the former can be further applied to making sense of the latter" (Traiger 2010: 52-3).[5]

On the whole, it appears that Vaihinger gleaned more from his early study of Hume regarding the nature of fictions than he readily admits in the preface to his *The Philosophy of As If*. At the very least, it is clear that Vaihinger's claim about Hume's understanding of fiction is mistaken. In fact, in my view, Hume's extensive examination of cognitive fictions – though largely obscure and unsystematic – is comprehensive, and often reaches similar conclusions to Vaihinger. In this discussion, I aim to redress Hume's absence from the history of fictionalism, as recapitulated by Vaihinger. To do so, I first provide a brief overview of Vaihinger's primary philosophical distinction, namely, that fictions may be divided into two kinds: (1) real fictions and (2) semi-fictions. I then illustrate how various Humean fictions fit the criteria of each kind. Subsequently, I argue that Vaihinger's philosophy follows Hume in several other related areas: (a) cognitive fictions are identified as remedies for psychological instability, (b) fictions are tied to contradictions embedded in human nature, and (c) fictions are reified as dogmas over time. Finally, I assess the aims of Vaihinger's critical positivism against Hume's mitigated skepticism. As a consequence, I hope to show that Hume's historical contribution to contemporary fictionalist discourse is significant and ought to be regarded as such.

1. *Hypothetical and Self-Contradictory Fictions*

I now turn to demonstrating that Hume's theory of fiction anticipates several of Vaihinger's arguments in *The Philosophy of As If*. Although Hume does not classify or define his use of fiction in any systematic or technical sense, a careful reading suggests that Humean fictions share several isomorphic features with Vaihingerian fictions. In this section, I begin by highlighting the

[4] Vaihinger subsequently seems to contradict himself when he makes the claim: "it is the particular merit of Kant to have shown that most ideational constructs are purely subjective. That they are fictions in our sense, i.e. fictions as a means for attaining certain purposes, he no more realised than Hume" (*PAI*: 107).

[5] For a complementary discussion of Hobbes' application of legal fictions to natural philosophy and politics, see Foisneau 2010.

main contours of Vaihinger's theory of fiction, and thereafter I discuss how Hume's fictions may be understood in light of Vaihinger's twofold distinction.

In Chapter XIX, Vaihinger unpacks his foundational division between two kinds of fiction: (1) Real Fictions (Self-Contradictory) and (2) Semi-Fictions (Hypothetical). Real Fictions are fictions identified by an internal contradiction. They serve the mind in so far as they act as expedients for the operation of thought. Once their service is complete, they drop out like the middle term of a syllogism. As Vaihinger argues, "the discarding of true fictions in *the course of a given mental operation* follows necessarily from their contradictory character – for, after all, our aim is to obtain non-contradictory results" (*PAI*: 98).

Real fictions are essentially artificial. They blend the given with the unthinkable. That is to say, they assume the impossible and make reality incomprehensible, in order to make it comprehensible. While that may seem paradoxical, real fictions behave in contradiction of the facts and interpolate impossible elements for reality such that they, in fact, make it more complicated than it appears to be. Some examples may serve to illustrate the thought. For Vaihinger, the Absolute and the Infinite are both examples of real fictions: he argues that "self-contradiction discloses itself particularly in the antinomies to which they give rise (cf. Kant's antinomies of the infinite, by means of which he proved that the idea of infinite space was subjective, or, in our terminology, fictional)" (*PAI*: 97). Real fictions, in other words, are most easily recognized by internal contradictions or antinomies.

Semi-Fictions, on the other hand, are a second kind of fiction in Vaihinger's taxonomy. These fictions are historically provisional and disappear in the course of time. Instead of assuming the impossible, semi-fictions assume the unreal. In other words, they are hypotheses. In attempting to discover the truth, they deviate or falsify reality as simpler than it appears to be. Semi-fictions establish natural laws, while real fictions act as scaffolding to be demolished after their intended use: "to the *verification* of the hypothesis corresponds the *justification* of the fiction" (*PAI*: 88-9). Vaihinger, quoting Hermann Lotze, argues "'every hypothesis claims to be not only a figure of thought, or a means of making thought concrete, but a statement of fact.' 'Everyone who sets up a hypothesis believes that he has extended the series of real facts by a happy divination of facts not less real though falling outside the range of his observation'" (*PAI*: 90). Thus, hypotheses seem to reify future contingents by virtue of the epistemic attitude involved.[6] Let us now

[6] For Hume, what I call 'reifying future contingents' refers to the process of completing the union or adding 'new' relations. Timothy Costelloe discusses a similar process of 'reification' and its relationship to metaphysical fictions (2018: 92-3). As Hume remarks, "when we turn our thought to a future object, our fancy flows along the stream of time, and arrives at the object by an order, which seems most natural, passing always from one point of time to that which is immediately posterior to it.

consider both types of fiction in, what I believe to be, their Humean manifestations.

2. *Hume's Real Fictions*

In the *Treatise*, Hume did not classify his use of fiction in an unambiguous manner. The lack of conceptual clarity poses an interpretive difficulty: are there different *types* of Humean fictions? While an exhaustive typology is outside the scope of the present discussion, I suggest that several Humean fictions anticipate Vaihinger's twofold classification. I first focus my discussion on real fictions. I argue that Hume's construal of (1) the fiction of identity and (2) fictitious unity represents two paradigm cases of Vaihinger's real fictions.[7]

The essential quality of real fictions is that they possess an internal contradiction. In Hume's investigation of the origin and content of several natural and philosophical relations, he discovers that both identity and unity (of a specific sort) are self-contradictory.[8] Consider Hume's analysis of fictitious unity first, namely, that "twenty men *may be consider'd as an unite*. The whole globe of the earth, nay the whole universe *may be consider'd as an unite*. That term

This *easy* progression of ideas favours the imagination, and makes it conceive its object in a stronger and fuller light" (*T* 2.3.7.8, SBN 430-1). The stronger and fuller light produces a stronger belief in certain hypotheses such that, in Costelloe's terms, we are apt to confer on them "a real existence they do not possess" (2018: 93). I pursue a more detailed discussion of the matter in the next section.

[7] It might be doubted here whether Hume considers identity a fiction. On my reading, Hume clearly suggests that it is: if identity is a *relation* of an object through time, then it must be a *fictitious* relation in so far as it is relating two incompatible ideas, namely, unity and number. Moreover, when Hume says that "[t]his fiction of the imagination almost universally takes place," he is referring to fictitious duration, which is the incompatible *relation* or *union* of an unchangeable object and changing perceptions (*T* 1.4.2.29, SBN 201). Thus, fictitious duration is synonymous with the relation of identity. Otherwise, what *exactly* distinguishes fictitious duration from identity? Finally, that identity is classified as a 'fiction' is well-supported in the literature; for scholars who refer explicitly to Hume's "fiction of identity," see Coventry 2007: 159; Landy 2018: 248; Waxman 1994: 206; Green 1968: 258; Ross 1991: 352; Traiger 1987: 389.

[8] Norman Kemp Smith, for instance, interprets the self-contradictory nature of these fictions in the following passage: "Hume is…insistent that the idea of identity, on examination, turns out to be a fiction, and so, like all fictions, to be due to the imagination. In employing this idea we profess to be travelling upon a path between unity and number, as impossible a path as any between existence and non-existence. We both do and do not assert unity; that is to say, we refuse to go to the length of number or diversity, and yet restrain ourselves from asserting a strict and absolute unity. *Every alleged instance of such identity is an illustration of this self-contradictory procedure*; a body is, we believe, both diverse and a unity, a self we believe to be individual and yet also complex, the same with itself and yet in never-ceasing change. For imagination, and therefore belief, there is no difficulty. Nature, in and through our natural beliefs, imposes the fiction upon us; and this notwithstanding its having no sanction in the data of sense, and though the problems which it raises are irresolvable for the understanding and reason" (Kemp Smith 1941: 475-6, italics added).

of unity is merely a fictitious denomination, which the mind may apply to any quantity of objects it collects together" (*T* 1.2.2.3, SBN 30). When the mind unites the idea of unity with any collection (i.e., number), it creates a self-contradictory fiction. Unity and number cannot both describe the same thing in the same respect; and yet, when twenty men are taken as a *single* unit, we seem to generate a self-contradictory fiction.

Keep in mind that I am specifically referring to a psychological contradiction, not a logical contradiction consisting of propositions. Real fictions, in light of Hume's empirical commitment to the Copy Principle, are contradictory in so far as a single fiction is the relational *union* of incompatible ideas. Indeed, it is crucial to remember that Hume's logic is predicated on his theory of ideas, in which each atomistic idea is derived from an impression of sensation or reflection. Therefore, when the *idea* of number and the *idea* of unity are united in the imagination, it forms a self-contradictory union of individuated ideas. Humean fictions are thus *relations* and not to be understood as ideas themselves. Fictions are not resolvable into simple impressions because they are imaginative constructions; they are not perceived via sensation or reflection. Nevertheless, is it plausible to say that Hume's concept of fictitious unity is a 'Real Fiction' in Vaihinger's sense? I consider three criteria in the following examination: (a) is there an internal contradiction? (b) is it an expedient for the operation of thought? (c) does the fictional aspect drop out in the course of reasoning?

First, the attribution of an internal contradiction to Hume's rendering of fictitious unity seems uncontroversial. Consider that "in our most familiar way of thinking," we suppose that "a thing is in a certain place, and yet is not there" or, in other words, "*of totum in toto & totum in qualibet parte*" (*T* 1.4.5.13, SBN 238). Naturally, our imagination unites, for example, ideas of extension with ideas of taste, where "upon reflection we must observe in this union something altogether unintelligible and contradictory" (*T* 1.4.5.13, SBN 238). In this respect, Hume subsequently refers to the symmetry between identity and unity:

> What I have said concerning the first origin and uncertainty of our notion of identity, as apply'd to the human mind, may be extended with little or no variation to that of *simplicity*. An object, whose different co-existent parts are bound together by a close relation, operates upon the imagination after much the same manner as one perfectly simple and indivisible, and requires not a much greater stretch of thought in order to its conception. From this similarity of operation we attribute a simplicity to it, and feign a principle of union as the support of this simplicity, and the center of all the different parts and qualities of the object (*T* 1.4.6.22, SBN 263).

Note that a feigned simplicity is united with the ideas of different parts to produce the fiction of unity. The unavoidable imaginative process is fundamental to human nature; real fictions are constitutive of the way in which humans perceive (via our senses) and generate (via our imagination) the world around us. For Hume:

> [O]ur ideas of bodies are nothing but collections form'd by the mind of the ideas of the several distinct sensible qualities… however these qualities may in themselves be entirely distinct, 'tis certain we commonly regard the compound, which they form, as One thing, and as continuing the Same under very considerable alterations. The acknowledg'd composition is evidently contrary to this suppos'd *simplicity*, and the variation to the *identity* (*T* 1.4.3.2, SBN 219).

In the case of both identity and unity, self-contradiction is involved. In the former case, there is a contradiction between sameness and alteration; in the latter case, there is a contradiction between one and the many.[9] For these reasons, I take it that fictitious unity satisfies the first criterion of Vaihingerian real fictions.

Second, to say Hume's concept of fictitious unity serves as an 'expedient for the operation of thought' ought to be granted. It is useful to think in aggregates, despite the fact that, on Hume's picture, aggregates are associated by the imagination, mind-dependent, and empirically unverifiable. That nations or sports teams are fictional unities, for instance, does not affect their utility. The perspective that, say, a chair is made up of smaller particles does not alter the idea of the chair as a unity; instead, it provides an alternative, though contradictory, perspective of the chair. It is *one* and *many* at the same time and in the same respect. Remember that 'in the same respect' is not to be understood in Aristotelian terms, but in terms of Hume's theory of ideas.

Finally, when we think in aggregates, the contradictory or fictional nature of the aggregate drops out of our reasoning. If we take 'England,' for instance, we may use it in our reasoning without acknowledging – or even being aware of – its self-contradictory *a priori* structure. Somehow, despite their logical incomprehensibility, self-contradictory fictions perform an essential role in the operation of thought. In referring to a nation, for instance, we may employ the fiction in our reasoning and arrive at truth, despite its fictionality. England, say, is above Spain (a true proposition) even though the relata involved (England and Spain, or the map, for that matter) are fictional aggregates of spatial unities.

[9] Observe that in *T* 1.4.3.2, Hume specifically says 'contrary' as opposed to 'contradictory.' In his explicit definition of contrariety, however, Hume seems to think that contrary ideas do form a kind of contradiction: "no two ideas are in themselves contrary, except those of existence and non-existence" (*T* 1.1.5.8, SBN 15). The exact definition of contradiction for Hume, however, presents a number of interpretive difficulties.

For Vaihinger, the ability to attain consistency via contradictory conceptual constructs is what defines real fictions. Logic is therefore a technology crafted for the *art* of thought. In the cognitive sphere, Hume appears to uncover an analogous distinction, such that some of our unions of ideas are plainly fictional – e.g., winged horses or fiery dragons – but other fictions are constructed as tools for pragmatic purposes (e.g., within logic, mathematics, and language). Identity and unity are fictional in the latter sense, with the important caveat that they are primarily constructed naturally and unavoidably by the imagination. It is not that humans have intentionally created these logical tools; rather, they emerge in our cognitive development – for, as Vaihinger reminds us, "[i]t must be remembered that the object of the world of ideas as a whole is not the portrayal of reality – this would be an utterly impossible task – but rather to provide us with an *instrument for finding our way about more easily in this world*" (*PAI*: 15).

The obvious parallel here is 'object permanence'. The concept drawn from contemporary psychology shares important features with Hume's fiction of continued existence. Object permanence also seems to involve the fiction of identity. That is, before some identical object can exist across time, there first needs to be a fictitious union of identity. Identity is thus prior in our cognitive development, and its self-contradictory nature drops out to yield a non-contradictory result: the vulgar supposition of enduring objects. Indeed, Hume directly tells us that identity is contradictory, but the vulgar supposition is non-contradictory (*T* 1.4.2.40, SBN 208). On the ground of the non-contradictory supposition, a new self-contradictory fiction arises: the *belief* in the feigned continued existence of all sensible objects (*T* 1.4.2.42, SBN 209).[10]

In Vaihinger's terminology, Hume's fictitious unities may be understood to blend the unthinkable with the given. The fiction of unity is *unthinkable* in so far as it is contradictory, but the constituents of the unity are *given* in our experience in the form of simple or complex ideas and impressions. It is thus an act of assuming the impossible (a contradiction) in order to make reality comprehensible, since, in these particular cases, the comprehension of things (other than simple unities) is dependent upon grouping together particulars for various reasons. For instance, we make sense of a basket of fruit – as a monadic

[10] Robert McRae, for instance, argues that fictitious duration (which is used to explain Hume's fiction of identity) has extensive implications in the *Treatise*, namely, that "out of this fiction are generated in a logically ordered series the basic metaphysical categories in terms of which the mind thinks, and all of them are fictitious" (1980: 124). John Passmore's claims that: "By the application of the same methods we gradually construct an entire system of fictions... The system rests on nothing more solid than 'trivial propensities of the imagination': our tendency to 'overcome' contradictions by constructing imaginary entities" (1968: 71).

collection – by collecting together various fruit in the basket and applying the fictitious denomination 'fruit' to them. Notice here the similarity between fictitious unity and Hume's account of abstract ideas. Although Hume's theory of general terms denies legitimate abstract ideas, it might be plausible to think that, in the specific case of abstract ideas being taken *as if* they represent more than one particular idea, the same analysis of fiction applies.[11]

An objection to the preceding account might seek recourse in Hume's rejection of contradictions, namely, "'tis in vain to search for a contradiction in any thing that is distinctly conceiv'd by the mind. Did it imply any contradiction, 'tis impossible it cou'd ever be conceiv'd" (*T* 1.2.4.11, SBN 43). There are two approaches that may obviate the difficulty. First, Hume's distinction between *conceiving* and *supposing* might indicate that real fictions are not in fact conceived; instead, they are only supposed. I return to this possibility in the next section. Second, it might be thought that Hume's perspectivism resolves the concern. That is, while unity and number form a self-contradictory union, only one side of the union may be conceived at any one time – even though the contradiction is always present upon reflection. Thus, the solution might be to treat identity and unity as fictions capable of being viewed under different, incompatible lights.

Robert Fogelin, for example, argues that Hume is committed to a kind of doxastic perspectivism. Specifically, he says that "Hume attempts to explain th[e] fiction of an unchanging object existing in time by using one of his favorite devices: placing things in *different lights*, or viewing them from *different perspectives*" (2009: 71). Or, in another example, "[t]he fiction of identity over time is the result of flip-flopping back and forth between these two perspectives [of unity and number]" (2009: 73). Indeed, as Hume remarks, identity "is an idea, which is a medium betwixt unity and number; or more properly speaking, is either of them, according to the view, in which we take it" (*T* 1.4.2.29, SBN 201). In my view, the same perspectivism that Hume appeals to in his analysis of identity may be applied to fictitious unity. Fictitious unity is both simplicity or oneness and multiplicity or number at the same time in the same respect *as* a single and individuated fiction. Although the contradiction is inconceivable or unthinkable in one sense, it is clearly explicable in another. The paradox might be construed like this: we are able to understand or explain what a contradiction is, at the same time we are somehow incapable of legitimately cognizing a contradiction. Therefore, how can we understand something that we cannot actually think? How, in other words, do we *know*

[11] Hume even deploys Vaihinger's 'as if' formulation in his discussion of abstract ideas: "The image in the mind is only that of a particular object, tho' the application of it in our reasoning be the same, *as if it were universal*" (*T* 1.1.7.6, SBN 20, italics added).

contradictions exist at all if we have never conceived them? Hume and Vaihinger seem to think that, despite their inherent unintelligibility, fictions are instrumental to the cognitive operations within human nature.[12]

For Hume, more broadly, while reason, or the embodiment of reason in the form of a philosopher, might find these conclusions repugnant, the vulgar pays them no mind – we will continue to think and act according to the principles of human nature – contradictory or not. Skeptics and rationalists alike are apt to resist the thought that contradictions are embedded in our cognitive architecture, but no matter the skeptical challenges against contradictions or the rational attempts to explain away contradictions, human nature is stronger than the normative force of reason. The authority of reason may indeed motivate the philosopher to root out all self-contradictory fictions wherever she or he may find them – but to no avail: the priorities of reason are subordinate to human nature's pragmatic concerns.

Again, in the *Treatise*, self-contradictory fictions ought to be understood as psychological phenomena. The same is true of Vaihinger's concept of fictions: "The fictive activity of the mind is an expression of the fundamental psychical forces; *fictions* are *mental structures*. The psyche weaves this aid to thought out of itself; for the mind is inventive; under the compulsion of necessity, stimulated by the outer world, it discovers the store of contrivances that lie hidden within itself" (*PAI*: 12).[13] Fictions, such as object permanence, are thus generated in virtue of circumstantial necessity.[14]

To conclude, I briefly review the three criteria of real fictions in so far as they pertain to the fiction of identity. Identity, as mentioned, involves, at the same time and in the same respect, both unity and number, such that it contains an

[12] For example, Hume reasons upon several ideas or principles he construes as unintelligible: (a) "the uniting principle among internal impressions" (*T* 1.3.14.29, SBN 169); (b) the "unintelligible instinct in our souls" (*T* 1.3.16.9, SBN 179); (c) in our familiar way of thinking, we employ 'confus'd notions' that, upon reflection, are observed to be "unintelligible and contradictory" (*T* 1.4.5.13, SBN 238). Unintelligibility, therefore, does not preclude a certain kind of understanding or utility.

[13] Compare Hume: "Mankind is an inventive species; and where an invention is obvious and absolutely necessary, it may as properly be said to be natural as any thing that proceeds immediately from original principles, without the intervention of thought or reflection" (*T* 3.2.1.19, SBN 484).

[14] While outside the scope of this paper, the same analysis may apply to Hume's discussion of justice, property, promises, and government in Book III of the *Treatise* in so far as these artifices are constructed for the stability of social circumstances. For an excellent comparison of epistemic fictions and social artifices in Hume's *Treatise*, see Sokolowski 1968. In G. R. S. Mead's interpretation of Vaihinger, logic and ethics are tied together in the following way: "in the beginning the natural man knows neither logical contradictions nor ethical conflicts; only in the course of evolution do these logical and ethical struggles arise out of the ground of the soul itself. And yet it is only in this strife that progress lies, so that the idea, or rather feeling, of sin is as much the principle of ethical improvement as contradiction the motive of logical perfection" (1913: 263).

internal contradiction. The fiction of identity is an expedient to thought in several domains, including logic and mathematics (the law of identity), morality and politics (personal identity), science (taxonomical identity or object identity), language (terminological identity), etc. The fictional or artificial aspect of identity, however, drops out in all of these domains. That is to say, there is no need to acknowledge the constructed or contradictory nature of any given identity relation, since reasoning upon these fictions provides genuine insight into the world. That a class or genus of animals is *fictional* in one sense, for instance, does not imply that it is arbitrary or fantastical. Similarly, object permanence may be fictional in the sense that our cognitive mechanisms generate it, but that does not imply that it is arbitrary or unreal.

3. *Hume's Semi-Fictions*

The second type of fiction in Vaihinger's account is the semi-fiction or hypothesis. This type of fiction is essential to scientific inquiry, especially theoretical science. In the course of science, we must treat hypotheses *as if* they are true in order to develop theoretical systems. But simply because a hypothesis turns out true does not make the hypothesis itself true. The problem of the ontological status of hypotheses in Hume's philosophy is particularly troubling, for, on his view, we seem to only have access to that which we have already experienced, namely, impressions or ideas. For instance, how might hypotheses be *copied* from impressions if hypotheses have not been perceived? Indeed, it appears contradictory to say that hypotheses about the *future*, are 'correspondent' and 'exactly represent' *past* impressions. Consider Hume's admonition:

> We must endeavour to render all our principles as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the simplest and fewest causes, 'tis still certain we cannot go beyond experience; and *any hypothesis, that pretends to discover the ultimate original qualities of human nature, ought at first to be rejected as presumptuous and chimerical* (*T* 0.8, SBN xvii, italics added).[15]

The difficulty in interpreting this passage is that hypotheses *by definition* go beyond experience.[16] Therefore, if we cannot go beyond experience, what kind of ontological or epistemic status attends hypotheses? Though Hume's use of

[15] Consider two additional principles: "That there is nothing in any object, consider'd in itself, which can afford us a reason for drawing a conclusion beyond it"; and, "that even after the observation of the frequent or constant conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience" (*T* 1.3.12.20, SBN 139).

[16] For Hobbes: "The *present* only has a being in nature; things *past* have a being in the memory only; but things *to come* have no being at all, the *future* being but a fiction of the mind" (1994: 14).

hypothesis is not univocal, I define Humean hypothetical fictions specifically as fictions that go beyond past relations or constant conjunctions to imagine or suppose 'new' constant conjunctions or relations; in other words, they are instances of 'completing of the union'.[17] Note that hypothetical fictions, on this interpretation, are not *ideas*, but relations. Thus, these kinds of hypotheses are not resolvable into simple impressions or ideas because they are imaginatively generated.

At first blush, there are several examples in the *Treatise* that may be characterized as hypothetical *fictions*. For instance, the "hideous hypothesis" of the immateriality of the soul or the universal substance in which all configurations of matter are supposed to inhere (*T* 1.4.5.19, SBN 241). Similarly, extension, conceived as perfectly simple and indivisible, is an untenable hypothesis (*T* 1.2.3.14, SBN 38). Another supposition that Hume explicitly designates a fiction is a kind of fictitious distance, in which "a body interpos'd betwixt two others may be suppos'd to be annihilated" (*T* 1.2.5.23, SBN 62-3). While the latter is most relevant here in so far as it is a vulgar hypothesis – i.e., a supposition that all of us hold at one time or another – there is a further hypothetical fiction that I will focus my discussion on: the vulgar supposition of enduring objects.

In this particular hypothetical fiction, "the vulgar *suppose* their perceptions to be their only objects, and at the same time *believe* the continu'd existence of matter" (*T* 1.4.2.43, SBN 209). Two claims are involved here: (a) the vulgar supposition and (b) the belief in the continued existence of matter. While I take the *belief* in the continued existence of objects to be a real fiction in the Vaihingerian sense, it is the vulgar *supposition* that, I submit, fulfills the criteria of the hypothetical fiction. Indeed, though it is likely that the belief in the continued existence of objects involves a contradiction so far as it involves the fiction of identity, the vulgar supposition seems to avoid it (though it is unclear whether it, in fact, succeeds). Still, Hume equivocates on the supposition/belief distinction, making it difficult to discern his exact view; for instance, he claims that "[t]he *supposition* of the continu'd existence of sensible objects or perceptions involves no contradiction" and that we have a "common *hypothesis* of the identity and continuance of our interrupted perceptions" (*T* 1.4.2.40, SBN 208, italics added; *T* 1.4.2.46, SBN 211, italics added).

To clarify Hume's position, then, I follow his distinction between supposition and belief. Therefore, the *supposition* is the common hypothesis of the

[17] I also include Humean 'suppositions' that meet this definition as hypothetical fictions. For Don Garrett: "To 'suppose' something, in Hume's usage, is to act in at least many respects as though one believed something, but without necessarily forming a lively idea of it. This also occurs, for example, when human beings and even animals suppose the uniformity of nature without formulating it as a belief" (2015: 52). Note Garrett's formulation of 'as though' in relation to Vaihinger's 'as if'.

vulgar involving no contradiction. The *belief* in the continued existence of objects, on the other hand, involves the contradiction of uniting object identity with the gaps among interrupted perceptions. It is the former that meets the definition of the hypothetical fiction, in as much as it goes beyond past relations or constant conjunctions to imagine or suppose 'new' constant conjunctions or relations. In other words, "we may easily indulge our inclination to that supposition. When [1] the exact resemblance of our perceptions makes us ascribe to them an identity, we may remove the seeming interruption by [2] feigning a continu'd being, which may fill those intervals, and preserve a perfect and entire identity to our perceptions" (*T* 1.4.2.40, SBN 208, numbers added). The first step is the supposition which goes beyond the exact resemblance relations in the memory to imagine a 'new' identity relation in order to complete the union; the second step is that a continued being is feigned to preserve the new relation and remove the interruption. Therefore, the supposition of a new relation leads the imagination to feign a continued being that confers existence on the relation. The former is a semi-fiction; the latter is a real fiction.

For Vaihinger, hypothetical semi-fictions intrinsically seek verification: "[t]he hypothesis has ultimately only a theoretical object, that of connecting facts and filling up the gaps in the connection, which experience shows to be numerous; and of establishing what is ultimately and primarily unalterable" (*PAI*: 88). Hume's description of generating the belief in continued existence seems to share this feature:

> [R]esemblance gives us a propension to consider these interrupted perceptions as the same; and also a propension to connect them by a continu'd existence, in order to justify this identity, and avoid the contradiction, in which the interrupted appearance of these perceptions seems necessarily to involve us. Here then we have a propensity to feign the continu'd existence of all sensible objects; and as this propensity arises from some lively impressions of the memory, it bestows a vivacity on that fiction; or in other words, makes us believe the continu'd existence of body (*T* 1.4.2.42, SBN 208-9). All sensible objects continue in existence via hypothetical fictions or suppositions that fill in unperceived gaps across interrupted perceptions.[18]

The repeated relations of resemblance prompt the imagination to complete the union by adding a new relation of identity, to which a feigned continued being is generated to confer existence on the relation. In other words, the sup-

[18] H.H. Price, in his *Hume's Theory of the External World* (1940), suggests that Hume may appeal to an 'As-if Theory' to possibly ground what he calls *unsensed sensibilia*. He believes, however, that such an appeal ultimately fails. It is unfortunate that Price does not genuinely engage with Vaihinger's actual work.

position "substitutes something conceived [or, in Hume's case, supposed] for what is actually given" and "avoid[s] difficulties by deviating from reality" (*PAI*: 80). By assuming the unreal, we may seek to verify the given, but whether the given is possible to conceive *without* fictional constructs is where Vaihinger and Hume must both face the ouroboros of self-reference. Wayne Waxman interprets Hume's self-referential paradox in this way:

> Hume's intention, in my view, was to conclude his analysis of human understanding by declaring his own theory of ideas – employed throughout the *Treatise* I to explicate relation, abstraction, space and time, necessary connection, and identity – to be itself a mere fiction of associative 'imagination, or the vivacity of our ideas' (*T*265). When the dragon thus swallows itself up from the tail, the stage is finally set for the skeptical denouement of *Treatise* I/iv/§7 (1994: 202).

Nevertheless, as Vaihinger notes, "the realization that imagination… plays a great part in science is one of the main advances of modern epistemology. In this respect Kant was quite correct and circumspect when he spoke of a 'transcendental imaginative faculty'" (*PAI*: 55). Subsequently, he adds the qualification that the discovery "of the whole conceptual world as mere products of the imaginative faculty was originally accomplished by Hume and Kant" (*PAI*: 63). Indeed, Hume's recognition of the 'empire of the imagination' within our cognitive landscape is a significant development in early modern philosophy. Unfortunately, due to the absence of his theory of fiction from the *Enquiries*, Hume's early and more radical views – those that specifically anticipate Vaihinger's philosophy of fictionalism – have generally been neglected.[19]

4. *Fiction and Psychological Stability*

That several of Hume's fictions fit the criteria of Vaihingerian fictions is only one part of the story. There are, in fact, several other areas where Hume's thought may have influenced Vaihinger, and therefore the history of philosophical fictionalism. Consider that Vaihinger and Hume both point to a psychological reason behind the mind's generation of fiction. Indeed, while Vaihinger does not mention Hume's appeal to psychological stability, he makes some strikingly complementary remarks. For instance, he argues that "[a]n idea that has once been accepted as objective has a stable equilibrium, the hypothesis an unstable one. The psyche tends to make every psychical content more stable and to extend this stability. The condition of unstable equilibrium

[19] For notable exceptions, see Costelloe 2018; Cottrell 2012; Hollinger 1977; Iser 1993; McRae 1980; Sokolowski 1968; Thielke 2003; Traiger 1987; Varzi 2013; Whelan 1985.

is as uncomfortable psychically as it is physically" (*PAI*: 125). The idea that the mind is more comfortable when its ideas are stable is a recurring theme in Hume's philosophy.[20] The imagination in particular appears to be manifestly hedonistic; it acts according to ease and pleasure.[21] Ideas that resemble each other make transitions *easier* and *smoother*, while completing unions of relations produces mental *pleasure*.

> Nothing is more certain from experience, than that any contradiction either to the sentiments or passions gives a sensible uneasiness, whether it proceeds from without or from within; from the opposition of external objects, or from the combat of internal principles. On the contrary, whatever strikes in with the natural propensities, and either externally forwards their satisfaction, or internally concurs with their movements, is sure to give a sensible pleasure (*T* 1.4.2.37, SBN 205-6).

The imagination naturally attempts to resolve uncertainty and instability among its ideas. To do so, it creates fictions. These solutions, however, turn out to be mere placebos – or, in Robert Fogelin's words, "an empty placeholder for a solution to a problem masquerading as a solution" (2009: 89). Instead of fictions resolving contradictions and uncertainties inherent in human nature, they disguise our cognitive infirmities. Thus, when philosophers open the hood to examine the mechanics of the mind, they notice contradictions built atop contradictions and hypotheses built atop hypotheses – in other words, a scene of pure terror for those committed to the dictates of reason and logical consistency. Some philosophers take the situation as incentive to construct even more fictions to escape the contradictions and unjustified suppositions they discover. But, for Hume, these fictions and contradictions are constitutive of how the mind operates. The task of the true skeptic is to accept the mind as it is – flaws and all – and humbly acknowledge our limitations.[22]

In *The Philosophy of As If*, the human mind designs fictions expressly for resolving psychical tension. Ironically, however, fictions cannot be discovered to *be* fictions, otherwise one of the main pragmatic reasons for their existence

[20] See Loeb 2002 for a detailed account of the relevance of stability in Hume's philosophy.

[21] I follow Timothy Costelloe's description here: "Hume identifies a hedonistic tendency that inclines the imagination always to seek and make an easy and smooth transition among ideas in order to form a union or complete a whole, from which it derives pleasure" (2018: 1).

[22] Dorothy Coleman makes the argument that: "True philosophy is sceptical concerning the natural illusions of the imagination just because it properly recognizes their illusory character. However, since detecting an illusion does not destroy it, even the experience of true philosophers continues to be shaped by the illusion. Consequently, true philosophers continue to experience a psychological opposition between natural beliefs about perceptions, even after the epistemic status of these beliefs has been determined. True philosophers free themselves from the psychological opposition between natural beliefs by mitigating their scepticism when engaged in practical activity" (1984: 150).

disappears. For instance, though the fiction of identity possesses an internal contradiction, the fiction drops out, making it seem as if the contradiction has disappeared. Vaihinger argues:

> If then we first compare the dogma with the hypothesis, we notice that the latter involves a condition of tension which must be exceedingly disagreeable to the mind. The mind has a tendency to bring all ideational contents into equilibrium and to establish an unbroken connection between them. An hypothesis is inimical to this tendency in so far as it involves the idea that it is not to be placed on an equality with the other objective ideas (*PAI*: 125).

Moreover, Vaihinger indicates that dogmas are more satisfying to the mind, whereas hypotheses and, even more so, fictions are disagreeable. The inequality among the statuses of our ideas is psychologically uncomfortable as opposed to having all of our ideas accepted as real and certain. The imagination then naturally constructs expedient fictions (that drop out) and connective hypotheses (that may be reified) to prop up a stable and real psychological experience. By undergoing reification or dropping out, fictions are like unseen stagehands moving around the wings of the mind, never once making an appearance on stage. It is only when fictions are no longer needed or when philosophers reflect on our mental processes that the grand production is revealed as artifice. And, while the philosopher's awareness of cognitive or perceptual contradictions might prove uncomfortable for *them*, for the vulgar, who are embedded in common life, fictions perform their role invisibly.

If all this is accurate, it stands to reason that philosophical schools of realism or idealism will inherently be more appealing than fictionalism for purely psychological reasons. Hume and Vaihinger, however, do not follow up that potentially rich line of reasoning. What we can say is that the two philosophers both think contradictions and unproven hypotheses are destabilizing to the mind. And yet, fictions are only successful if they are believed, the same as an audience must buy into the premise of a film or novel to experience its rewards.

That is perhaps the reason why Hume and Vaihinger alike provoke such strong reactions. The through line of our perceptual story is broken by their conclusions. To reveal that central concepts in human nature, like infinity and identity, are, so to speak, 'wearing no clothes,' is bound to stir up ire, especially among those who truly believe those concepts to be real. To claim that God is fictional among the devout or that identity is fictitious among the logicians is to commit doctrinal heresy. Though Hume and Vaihinger had noble aims in mind upon unveiling the artificiality of our conceptual structures, contemporaneous philosophers who nurtured dogmas of the day faced far less criticism. Indeed, Hume's early philosophy precluded him from ever receiving an academic post.

Vaihinger, on the other hand, waited more than thirty years to publish his seminal work, for he said it was not ripe at the time (Fine 1993).[23] While Vaihinger subsequently garnered a large following (likely affected by his already established fame as an interpreter of Kant), the logical positivists were hostile to his teaching. Others like H.L. Mencken criticized Vaihinger's philosophy as "not a system of philosophy, in any sense; it is simply a foot-note to all existing systems. Moreover, it is not a foot-note of much solid value. It is curious, but it is unimportant" (1924: 255).

Notably, for Vaihinger, the *main result* of his investigation "is that *contradiction* is the driving force of thought and that without it thought could not attain its goal at all; that it is immanent in discursive thought and is one of its constituent elements" (*PAI*: 108). Hume's natural fictions are likewise irresistible and unavoidable in human nature. Consider that Hume never ends up resolving many of the contradictions he discovers in human nature, let alone the manifest contradiction between causal reasoning and the continued existence of objects. Instead, he despairs, deciding to relax his bent of mind and engage in some amusing activity. Vaihinger, however, takes the cause of Hume's despair and revels in it, extolling it as the main discovery of his investigation. The contradictory nature of real fictions is the foundation of all discursive thought for Vaihinger, whether that be logic, science, mathematics, or religion. Contradictory fictions are natural to the operation of human thinking. In that sense, there is no reason to despair, for there is nothing inherently wrong with contradiction.

In this respect, Hume and Vaihinger received similar attacks by interpreters who mistook them for extreme skeptics – though neither philosopher accepted the attacks as valid.[24] On one hand, Hume's critics took his philosophy to be primarily skeptical and self-defeating. T.H. Green, for instance, claimed that Hume's "method, which began with professing to explain knowledge, showed knowledge to be impossible. Hume himself was perfectly cognisant of this result" (1997: vii). Vaihinger's critics attacked him on similar grounds. Morris R. Cohen bemoaned the fact that "since the publication of Vaihinger's *Philoso-*

[23] Note that Fine interprets Vaihinger's philosophy to exhibit "a strong British influence – especially due to Berkeley on the philosophy of mathematics and Hume on impressions and the imagination" (1993: 4).

[24] In particular, Vaihinger claims that the "term 'Scepticism' has occasionally been applied to the Philosophy of 'As if' and its systematic doctrines; but this is not correct, for scepticism implies a theory which raises doubt or questioning to the dignity of a principle. The Philosophy of 'As if', however, has never had a trace of this attitude" (*PAI*: xlii). Instead, he says that "'Relativism' would be more applicable to the Philosophy of 'As if', in so far as it denies all absolute points (in mathematics just as in metaphysics) and shows a natural affinity with the theory of relativity both of the past and the present" (*PAI*: xlii).

phie des Als Ob, there has been an increased general recognition of the importance of fiction as construction in science. But the subject has been beclouded by the monistic mania. By trying to show that everything is a mental construction, the distinction between fact and fiction is really obliterated" (1923: 484). Vaihinger, however, does not subscribe to such a view, since he announces that his investigation aims to introduce fiction as a third member in a system of logical science, alongside induction and deduction. According to Vaihinger:

Thought conducts us automatically to certain illusory concepts just as in vision there are certain unavoidable optical errors. If we recognise this logical illusion as necessary, if we accept the fictions established thereby with a full realisation of their significance and, at the same time, see through them (e.g. God, liberty, etc.) then we can cope with the logical resultant contradictions as necessary products of our thinking, by recognising that they are the inevitable consequences of the inner mechanism of thought itself (*PAI*: 133-4).

Contradiction therefore need not be seen as destructive to epistemic belief; it may be constitutive of epistemic belief. On such an account, ancient and modern philosophers, whose theories attempt to escape contradiction, are pursuing the impossible. It is only when we accept contradiction as part of human nature that we may formulate a proper science of man.[25] Thus,

true philosophy approaches nearer to the sentiments of the vulgar, than to those of a mistaken knowledge. 'Tis natural for men, in their common and careless way of thinking, to imagine they perceive a connexion betwixt such objects as they have constantly found united together… But philosophers… have sufficient force of genius to free them from the vulgar error… but not sufficient to keep them from ever seeking for this connexion in matter, or causes. Had they fallen upon the just conclusion, they wou'd have return'd back to the situation of the vulgar, and wou'd have regarded all these disquisitions with indolence and indifference. At present they seem to be in a very lamentable condition, and such as the poets have given us but a faint notion of in their descriptions of the punishment of *Sisyphus* and *Tantalus*. For what can be

[25] Manfred Kuehn's approach to Humean contradictions is influential here. Particularly, he argues that "Hume believed that he had, by means of what he called the 'experimental method of reasoning,' manifest contradictions that are symptomatic of conflicting principles of the mind. The contradictions should, therefore, not be excused or explained away, but they should be fully acknowledged. Hume needs no defense here. For he might actually have considered this discovery of the antinomical character of the human mind one of his most important achievements. In any case, I believe that Hume's metaphysics can be understood correctly only if we take into account very carefully all the consequences of the antinomical dimension of his thought" (1983: 36-7). In my interpretation, I resist using Kantian language, and instead use Hume's preferred term: contradiction.

imagin'd more tormenting, than to seek with eagerness, what for ever flies us; and seek for it in a place, where 'tis impossible it can ever exist? (*T* 1.4.3.9, SBN 222-3).

5. *Hume's True Philosophy and Vaihinger's True Criticism*

Vaihinger's penchant for legislating philosophical laws is conspicuous across his work. Most well-known is his 'Law of the Preponderance of the Means over the End.' Based on his reading of Schopenhauer – in which thought is originally a means to the will's end – he expands the idea to capture the tendency for an "original means working towards a definite end… to acquire independence and to become an end in itself" (*PAI*: xxx). Fictions and hypotheses, therefore, often originate as means to achieve some cognitive end, but after time, they transform into ends themselves. A once practical purpose is forgotten, leaving in its place a theoretical thought shorn of its origins.[26] Taking the thought as theoretically meaningful subsequently generates a number of insoluble, even unintelligible, questions – for instance, what is the square root of -1?

Less well-known, however, is Vaihinger's 'Law of Ideational Shifts'. This law describes three developmental stages, in which ideas either pass from fictions to hypotheses to dogmas or the reverse. Across history, Vaihinger cites various examples of the law at work. For instance, Newton and Leibniz' fictitious fluxions and differentials, which "became hypothetical entities, and later, dogmas. After that there was regression" (*PAI*: 131). Vaihinger adds to this a number of other examples from Platonic myths to Kant's categories to the social contract.[27] The reason for these shifts is, again, psychological tension. Take hypotheses: the only legitimate way to support a hypothesis is via probability or, what Vaihinger calls, 'repeated confirmation'. For some ideas, though, to achieve any sort of stabilizing probability may "involve centuries of labour [or] in many cases be quite impossible. So the psyche circumvents it by simply transforming the hypothesis into a dogma by illegitimate methods" (*PAI*: 125).

Unlike Comte's law of three stages – i.e., (1) Theological (fictitious), (2) Metaphysical (Abstract), and (3) Scientific (positive) – Vaihinger claims that his law "stresses the *formal* change of the ideas themselves, whose content remains

[26] The thought is reminiscent of Nietzsche's "On Truth and Lie in a Nonmoral Sense." Vaihinger dedicates his final chapter to an examination of Nietzsche's 'Doctrine of Conscious Illusion'. See *PAI*: 341-362.

[27] Vaihinger's interpretation of Kant is unorthodox in this respect, for, as he says, "[t]he law of ideational shifts can be very clearly demonstrated from the fate that befell the Kantian concepts among his disciples. The fictive device tends more and more to become a (false) hypothesis. The maintenance of the pure result is, after all, quite difficult; for man has an inclination towards dogmatism. The study of Hume and Comte has proved a corrective to the errors of Kant and his successors" (*PAI*: 153).

always the same, whereas according to Comte it changes" (*PAI*: 132). In other words, it is primarily an epistemic theory of fiction rather than a metaphysical theory. The same holds true for Hume. Traiger argues that "[a]lthough there are different fictions, Hume has a core notion of fiction which is fundamentally epistemological rather than ontological" (1987: 382). Fictions, as a completion of the union of relations, are formal configurations *of* ideas; thus, fictions are consistent with Hume's commitment to empiricism and rejection of abstract ideas.

Recall Hume's true philosophy, as expressed in the passage at the end of the previous section: notice there is another triadic shift with respect to epistemic judgment. In human nature, everyone begins with a vulgar conception of the world, a conception that, for most, remains unexamined. For the few who do venture into philosophical reflection, they soon discover that the vulgar conception is rife with contradictions and unverifiable fictions.[28] The experience is unsettling, ultimately motivating them to search for answers and resolutions. Satisfaction is soon found in the dogma of false philosophy, where most philosophers remain, unaware of the ideological cave that protects them. The true philosopher, however, ends up recognizing false philosophy for what it is, and thereafter descends back to vulgar condition, but now with knowledge that natural illusions are both cognitively irresistible and pragmatically useful despite their violation of reason's decrees. In Hume's words, there is "a gradation of three opinions, that rise above each other, according as the persons, who form them, acquire new degrees of reason and knowledge. These opinions are that of the vulgar, that of a false philosophy, and that of the true" (*T* 1.4.3.9, SBN 222).

The construal of identity in 'Of scepticism with regard to the senses' presents an archetypal case of both Vaihinger's Law of Ideational Shift and Hume's True Philosophy. At first, there is a fiction: in this case, the real fiction of identity. It contains an internal contradiction; unity and number are both present at the same time and in the same respect. The contradiction goes unnoticed by the vulgar, since, according to the view in which we take it, only one side of the conjunct is ever apparent. The identity relation is consequently applied to various objects, which gives rise to a new hypothesis: the vulgar supposition of enduring objects. The initial fiction then becomes a hypothesis by virtue of,

[28] Annette Baier claims that certain fictions are not false but "are unverifiable, and so unverified" (1991: 103). Baier further argues that they "are not empirically verifiable or falsifiable, so they are '*a priori*'… [A]nd this is quite different from calling them false. What is provably false is to deny that they are fictions, and Hume does think that we are prone to such falsehoods. Fictions are plausible stories we tell ourselves to organize our experience… Fictions structure our version of ourselves and our environment, making us and it 'real and durable'… Like the poets Hume discusses at T. 121-122, we start from what is familiar to us, our perceptions, and build from that a 'system' that goes beyond what we strictly know to be true" (1991: 103-4).

say, the law of ideational shift.[29] Third, and finally, the supposition is reified into a dogmatic belief in the continued existence of objects – that is, a real fiction taken to exist externally and mind-independently. Hence, the origin of how external objects came to be believed – and all the mental scaffolding used in the construction – drop out in order to achieve broader cognitive aims: psychological stability and pragmatic utility.

For the philosopher, the psychological stability is short-lived: upon philosophical reflection, the relation of identity is discovered to be self-contradictory, and the vulgar supposition in enduring objects is discovered to be mere hypothesis. As Vaihinger says, "when experience and reflection have gradually made these dogmas doubtful… [t]he psyche still tries to adhere to them… when this is no longer possible with a stable equilibrium, when the position has been too much shaken, then it contents itself with the unstable equilibrium of the hypothesis. The dogma becomes an hypothesis and the idea is reduced in value by one degree" (*PAI*: 127). The philosopher still does not let the matter go; instead with "sufficient force of genius to free them from the vulgar error… but not sufficient to keep them from ever seeking for this connexion in matter, or causes," the philosopher constructs a new fiction: the double existence of objects and perceptions (*T* 1.4.3.9, SBN 223). It is only "by passing thro' the common hypothesis of the identity and continuance of our interrupted perceptions" that we are led to this new opinion (*T* 1.4.2.46, SBN 211). The fiction of double existence, which also Hume calls a philosophical hypothesis, shares the same contradictory nature as the original fiction of identity; though, this time, it is the monstrous offspring of two contrary principles: that our resembling perceptions are continued and uninterrupted (via the imagination) and that our resembling perceptions are interrupted and distinct (via reason). Similar to the fiction of identity, we elude this new contradiction by constructing the fiction of double existence, which justifies and illusively renders the two contrary principles consistent (*T* 1.4.2.52, SBN 215).

The true philosopher must therefore rise above the false philosophy of double existence by realizing that she or he has simply been building fictions and hypotheses entirely in the air.[30] There is no recourse and no resolution, for "reason

[29] Note that Vaihinger is unclear on exactly how the law of ideational shift operates. In another place, he says that "[t]hought begins with slight initial deviations from reality (half-fictions) [i.e., semi-fictions], and, becoming bolder and bolder, ends by operating with constructs that are not only opposed to the facts but are self-contradictory" (*PAI*: 16). In this case, Hume's continued existence would be the initiating hypothesis that would then lead to the self-contradictory fiction of identity. Part of the reason for Vaihinger's equivocation on this point may be the continuous revisions he made to his theory over the course of his life.

[30] Compare *D* 10.30.

is incapable of dispelling these clouds" (*T* 1.4.7.9, SBN 269). Instead, the true philosopher returns "to live, and talk, and act like other people in the common affairs of life" (*T* 1.4.7.10, SBN 269). While philosophy may not be able to cure the mind of its contradictory fictions and hypotheses, it nonetheless serves a useful function. It enables us to recognize fictions *as* fictions. By doing so, it encourages "mild and moderate sentiments" and prevents dogma – indeed, philosophy is more harmless than theology in this very respect, since theological dogma and superstition "seizes more strongly on the mind, and is often able to disturb us in the conduct of our lives and actions" (*T* 1.4.7.13, SBN 271-2). Philosophical dogmas, on the other hand, are only ridiculous. Of course, here we have yet another contradiction in human nature: dogmas provide the mind with the most satisfaction and yet dogmatic belief poses the most danger. For Vaihinger:

> Dogmatism is a form of logical optimism which approaches the logical functions and their products with unbounded confidence, regards thought with an admiration and satisfaction so exaggerated that doubts are not raised at any point. The logical infallibility of thought is adhered to by the logical optimist as though it were a Gospel in which he blindly believed; and with the same intolerance that accompanies religious superstitions he regards the logical form in which he happens to think as better than any other. This logical optimism is harmless and innocent enough when found among primitive people, but it is a questionable attitude and becomes definitely dangerous and disastrous when encountered in men of a more advanced type (*PAI*: 162).

Though the final claim is perhaps the inverse of Hume's position, the prevention of dogmatic belief is the root of Vaihinger's concern here. Skepticism, on the other hand, is not dangerous but 'barren' for Vaihinger; and, for Hume, it is 'impossible' in the case of *antecedent* skepticism. Both philosophers, however, sanction skepticism in preparing the way to what Vaihinger calls the "critical attitude which we ought to adopt" (*PAI*: 163). Hume likewise prescribes excessive skepticism as a plausible path for dogmatists to realize their epistemic limitations, and thereby adopt a general attitude of moderate or mitigated skepticism (*E* 12.24, SBN 161-2). Vaihinger endorses a rough equivalent, a position he identifies as True Criticism or Logical Positivism.[31] The aim of his position is to navigate between skepticism, which he calls logical pessimism, and dogmatism, which he calls logical optimism. In so doing, we may examine our thought technologies in a dispassionate and unprejudiced manner.

[31] It is claimed that Vaihinger may have been the first to use the phrase 'logical positivism'; see Stoll 2020. Note that Vaihinger's use of logical positivism is considerably different from later usages. In his autobiographical introduction, Vaihinger alternatively titles his philosophy "Idealistic Positivism" or "Positivist Idealism" (*PAI*: xli). He also denominates his philosophical position as 'Critical Positivism'.

With logical pessimism it frees itself from childish beliefs in the power and unlimited validity of thought, and with optimism it holds firmly to the fact of the ultimate practical coincidence of thought and existence. The valuable outcome of pessimism is the habit of seeing in these conceptual constructs primarily nothing more than subjective products. Instead, therefore, of demanding with the dogmatist that we accept their reality until their unreality is proved – a thesis that from a practical point of view is the only useful one – it reverses the process and *mutatis mutandis* applies the juristic formula 'Quisque praesumatur malus, donec probetur bonus,' demanding that every logical product and every logical function be taken for what it actually is, a mere logical construct; and insisting on a special proof before the reality of any given mental construct or logical form is assumed. Theoretically this is the only valid and useful principle (*PAI*: 163).

To that end, true philosophy for Hume and Vaihinger is the recognition of the human mind as unavoidably embedded in nature and its processes. While Vaihinger understands logic as cognitive technology within a broader strategy for evolutionary success, Hume, without Darwin's influence, holds a strikingly complementary position. For Hume, "[t]he sole end of logic is to explain the principles and operations of our reasoning faculty, and the nature of our ideas" (*T* 0.5, SBN xv). Our ideas are ultimately derivative of impressions, but the source of our impressions we do not know and, presumably, "must be resolv'd into *original* qualities of human nature" (*T* 1.1.4.6, SBN 13). Our reasoning faculty is likewise tied to natural processes; indeed, as a mental colony subservient to the empire of the imagination, reason is fundamentally dependent upon the principle of the association of ideas. Consequently, logic cannot detach from the baseness of nature's operations; its purpose is to explain and discover fictions and hypotheses, not remove, or resolve them. Though it is within reason's jurisdiction to correct and modify our natural propensities via reflective practices, there remains a persistent danger of reason attempting to usurp authority by pretending to replace our natural and imaginative illusions with purported 'truths' of the understanding.

Vaihinger and Hume's mutual prescription to recognize fiction *as* fiction and artifice *as* artifice is applicable to the individual and the collective alike. In both cases, it requires the capacity to contend with the psychological or social tension that necessarily arises from such a recognition. Consider, for example, the Kuhnian paradigm shift: the dogma of normal science is disrupted by the observation of anomalies. Anomalies are akin to contradictions in so far as they are discrepant with established theories and lawlike regularities of science. The recognition of the anomalies leads to a crisis or tension, where scientists develop a range of competing alternatives. When one of the alternatives is select-

ed, a revolution occurs such that the entire scientific domain adopts the new dogma, much to the (psychological) satisfaction of the scientific community.

Philosophical theories appear to undergo similar revolutions in an ongoing attempt to account for contradictions and unverified hypotheses. New philosophical theories are accepted as 'progress,' despite the fact that, if the past resembles the future, another revolution is surely waiting around the corner. For Vaihinger and Hume, I suspect, many of the paradigm shifts that have taken place are merely the advent of false philosophies, the fate of which is to be scraped from our cultural palimpsest to make space for new dogma. For Vaihinger, that is not progress: progress is the collective acknowledgement of our natural condition, one that is governed by self-contradictory fictions and hypotheses at the core of human cognition. He puts it like this:

> Progress can be discerned not only in the logical conscience of mankind in that the contradictions in fictions are noticed, but also in the logical capacity. For to maintain a fiction as a fiction implies a highly developed logical mind, one that does not surrender too precipitately to the equilibratory impulse but carefully distinguishes between means and end. To maintain the purely critical standpoint as represented by Hume and Kant, considerable mental energy is required. All attempts subsequent to Kant are nothing but attempts, and very premature attempts, to resolve that condition of tension which though uncomfortable at least disturbs mental slumber (*PAI*: 132).

The challenge then in maintaining this kind of global fallibilism is that it requires a highly developed logical mind or theoretical domain. Whether on the individual or collective level, it is exceedingly difficult to walk the tightrope of the purely critical standpoint. But difficulty should not prevent the inquirer or inquiry from pursuing this epistemic aim. In recent decades, the proliferation of fictionalist theories suggest that Vaihinger's prescription is starting to be taken more seriously – even though the majority of philosophers still subscribe to various forms of realism.[32] At the very least, it is promising that the concept of fiction is being treated with more concern in contemporary philosophical literature. Indeed, as G. R. S. Mead said, 'fiction' is "perhaps the most provocative epithet in the vocabulary of philosophy" (1913: 264). And, whether or not his claim remains true, it nevertheless ought to remind us of a longstanding polemical strategy employed by philosophers in wielding fiction as a weapon – a tactic that, in the history of philosophy, traces back to Plato's self-serving demarcation between *muthos* and *logos*. In Vaihinger's systematic rendering of

[32] In Chalmers and Bourget's now-famous study, over 80% of philosophers either accept or lean toward realism about the external world, while over 55% either accept or lean toward moral realism. See Bourget *et al.* 2014.

philosophical fictionalism, prefigured by Hume's early philosophy, the character of fiction is recast as protagonist instead of villain, preparing the way for the current epistemic shift. Away from dogma and back to hypothesis, the hero of truth is thrown into doubt, and philosophers begin to wonder: what if 'truth' all along has simply been the most expedient type of error?

Brent Delaney
Yorkville University
brentdelaney@gmail.com

References

Baier, Annette C, 1991, *A Progress of Sentiments: Reflections on Hume's* Treatise, Harvard University Press, Cambridge.

Bourget, David *et al.*, 2014, "What Do Philosophers Believe?", in *Philosophical Studies*, 170: 465-500.

Cohen, Morris R., 1923, "On the Logic of Fiction", in *The Journal of Philosophy*, 20: 477-88.

Coleman, Dorothy P., 1984, "Hume's 'Dialectic'", in *Hume Studies*, 10: 139-55.

Costelloe, Timothy M., 2018, *The Imagination in Hume's Philosophy: The Canvas of the Mind*, Edinburgh University Press, Edinburgh.

Cottrell, Jonathan, 2012, *Hume on Fiction: Space, Time, Number, and the Mind*, Ph.D. diss., New York University.

—, 2016, "A Puzzle about Fictions in the *Treatise*", in *Journal of the History of Philosophy*, 54: 47-73.

Coventry, Angela, 2007, *Hume: A Guide for the Perplexed*, Continuum, London.

Fine, Arthur, 1993, "Fictionalism", in *Midwest Studies in Philosophy*, 18, 1: 1-18.

Fogelin, Robert J., 2009, *Hume's Skeptical Crisis*, Oxford University Press, New York.

Foisneau, Luc, 2010, "Elements of Fiction in Hobbes's System of Philosophy", in Richard Scholar - Alexis Tadié eds., *Fiction and the Frontiers of Knowledge in Europe, 1500-1800*, Ashgate Publishing Company, Burlington.

Garrett, Don, 2015, *Hume*, Routledge, London - New York.

Green, T.H., 1968, *Hume and Locke*, ed. Ramon M. Lemos, Crowell, New York.

—, 1997, ed. Peter P. Nicholson, *Collected Works of T.H. Green*, Thoemmes Continuum, London.

Hobbes, Thomas, 1994, ed. Edwin Curley, *Leviathan*, Hackett Publishing Company, Indianapolis.

Hollinger, Robert, 1977, "Two Kinds of Fictionalism", in *The Monist*, 60: 556-67.

Hume, David, 1975, ed. L. A. Selby-Bigge, 3rd ed. P. H. Nidditch, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, Clarendon Press, Oxford.

Hume, David, 1978, ed. L.A. Selby-Bigge, 2nd ed. by P. H. Nidditch, *A Treatise of Human Nature*, Clarendon Press, Oxford.

—, 2007, ed. Peter Millican, *An Enquiry concerning Human Understanding*, Oxford University Press, Oxford. [*E*]

—, 2007, eds. David Fate Norton - Mary J. Norton, *A Treatise of Human Nature, Vol. 1: Texts*, Clarendon Press, Oxford. [*T*]

—, 2007, ed. Dorothy Coleman, *Dialogues concerning Natural Religion and Other Writings*, Cambridge University Press, Cambridge. [*D*]

Iser, Wolfgang, 1993, *The Fictive and the Imaginary*, The Johns Hopkins University Press, Baltimore.

Kemp Smith, Norman, 1941, *The Philosophy of David Hume: A Critical Study of Its Origins and Central Doctrines*, Palgrave Macmillan, New York.

Kuehn, Manfred, 1983, "Hume's Antinomies", in *Hume Studies*, 9: 25-45.

Landy, David, 2018, *Hume's Science of Human Nature: Scientific Realism, Reason, and Substantial Explanation*, Routledge, New York.

Loeb, Louis E., 2002, *Stability and Justification in Hume's* Treatise, Oxford University Press, New York.

McRae, Robert, 1980, "The Import of Hume's Theory of Time", in *Hume Studies*, 6: 119-32.

Mead, G.R.S., 1913, *Quests Old and New*, G. Bell & Sons, London.

Mencken, H.L, 1924, "Philosophers as Liars", in *The American Mercury*, October: 253-55.

Nietzsche, Friedrich, 2010, "On Truth and Lie in a Nonmoral Sense", in Taylor Carman, trans. and ed. *On Truth and Untruth: Selected Writings*, HarperCollins, New York.

Passmore, John, 1968, *Hume's Intentions*, Gerald Duckworth & Co., London.

Price, H.H., 1940, *Hume's Theory of the External World*, Clarendon Press, Oxford.

Ross, Don, 1991, "Hume, Resemblance and the Foundations of Psychology", in *History of Philosophy Quarterly*, 8: 343-356.

Russell, Paul, 2008, *The Riddle of Hume's* Treatise*: Skepticism, Naturalism, and Irreligion*, Oxford University Press, New York.

Sokolowski, Robert, 1968, "Fiction and Illusion in David Hume's Philosophy", in *The Modern Schoolmen*, 45: 189-225.

Stoll, Timothy, *Hans Vaihinger*, ed. Edward N. Zalta, The Stanford Encyclopedia of Philosophy, 2020, https://plato.stanford.edu/archives/spr2020/entries/vaihinger/.

Thielke, Peter, 2003, "Hume, Kant, and the Sea of Illusion", in *Hume Studies*, 29: 63-88.

Traiger, Saul, 1987, "Impressions, Ideas, and Fictions", in *Hume Studies*, 13: 381-99.

—, 2010, "Experience and Testimony in Hume's Philosophy", in *Episteme*, 7: 42-57.

Vaihinger, Hans, 1935, *The Philosophy of 'As if': A System of the Theoretical, Practical and Religious Fictions of Mankind*, trans. by C.K. Ogden, Second Edition, Kegan

Paul, Trench, Trubner & Co., & Harcourt, Brace & Company, London & New York. [*PAI*]

Varzi, Achille, 2013, "Fictionalism in Ontology", in C. Barbero, M. Ferraris and A. Voltolini eds., *From Fictionalism to Realism*, Cambridge Scholars Publishing, Newcastle.

Waxman, Wayne, 1994, *Hume's theory of consciousness*, Cambridge University Press, Cambridge.

Whelan, Frederick G., 1985, *Order and Artifice in Hume's Political Philosophy*, Princeton University Press, Princeton.

Consciousness lived through. Husserl's reception of Brentano's inner consciousness

Filippo Nobili

Abstract: The paper aims to examine Brentano's account of inner consciousness and to assess its reception in Husserl's early works. Starting from the preliminary definition of psychic phenomena and some basic distinctions such as those between inner perception and observation, primary and secondary object, etc., I discuss Brentano's later thought in the light of his theory of relation and temporality, revealing a certain inconsistency with his initial assumptions. Subsequently, I examine Husserl's critical reception of inner consciousness in the *Logical Investigations* (1901) and in his lectures up to 1904-05, that is, up to the first in-depth thematization of temporality, to which inner consciousness will be inextricably related. Indeed, Husserl's reworking of the inherited psychologistic lexicon helps to trace a *prehistory of his phenomenology of time* and to better understand the paradigmatic detachment of phenomenology from descriptive psychology.

Keywords: Brentano, Husserl, inner consciousness, inner perception.

As it is well known, Husserl's thematization of the question of time in his lectures and manuscripts focuses on the *innere Zeitbewußtsein*. In this sense, he roots his analyses in the notion of inner consciousness inherited from Brentano's psychology. A peculiar blending of synthesis constituting time and self-awareness results from this theoretical choice by Husserl, triggering a major, albeit subterranean, shift in transcendental phenomenology.[1] This paper aims to clarify Brentano's account of inner consciousness and Husserl's critical reception of that crucial theme in some of his early works.[2] In so doing, a prehistory of Husserl's reflections *On the Phenomenology of the Consciousness of Internal Time*

[1] See my Nobili 2022 for a comprehensive account of the impact of Husserl's analyses of time-consciousness on his understanding of phenomenological idealism.

[2] For reasons of space, I will limit myself to the first edition of the *Logical Investigations* (1900-01) – in which Husserl's preliminary works about logic and psychology are improved – and to the subsequent lectures on the *Allgemeine Erkenntnistheorie* (*WS* 1902-03) and the *Hauptstücke aus der Phänomenologie und Theorie der Erkenntnis* (*WS* 1904-05), where Husserl's detachment from Brentano becomes more definite. I will not consider the *Logik-Vorlesungen* (1902-03) as hardly relevant to the topic at hand.

philinq XII, 1-2024, pp. 37-60
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.534
Submitted: 2 June 2024
Accepted: 31 August 2024

will be traced, elucidating one of its main prerequisites. Moreover, starting from some preliminary terminological remarks in the *Logical Investigations*, the paradigmatic departure of phenomenology from descriptive psychology will progressively emerge on the basis of Husserl's reassessment of inner consciousness, that is through the transvaluation of Brentano's access key to psychic phenomena.

1. *Brentano on inner consciousness*

1.1 *Perception, observation and judgement*

Since its first formulation in 1874, Brentano's descriptive psychology sought its executive pivot in the inner perception (*innere Wahrnehmung*). On the possibilities and limits of functioning of the latter, Brentano based his philosophical project of clarifying the unspoken assumptions of psychic life, and thus of every other branch of human knowledge related to it. Inner perception was regarded as "the primary source of the experiences essential to psychological investigations" (Brentano 1973: 48; Eng. tr. 26). In the failure to characterize it, that is, the inability to distinguish inner perception from the phenomenon of inner observation (*innere Beobachtung*), lies the risk of rejection of psychological research *tout court* by young scholars.[3]

Even before being a basic heuristic tool for promoting a privileged access to the psychic dimension, inner perception is already employed to define psychic phenomena as such. In fact, Brentano recalls how the latter is characterized by: (1) being or being based "*upon presentations*," (2) lacking, unlike physical phenomena, the feature of "*extension*," (3) the peculiar trait of "*intentional in/existence*," (4) being the "exclusive *object of inner perception*," (5) appearing to us, "in spite of all their multiplicity, *always* […] *as a unity*" (136 f.; 74 f.). Let us develop point (4).

"We call – Brentano says – the perception of a mental activity which is actually present in us 'inner perception'", and therefore "the consciousness which is directed upon it 'inner consciousness'" (141 n.; 78 n.). The emphasis placed on the *innere Bewußtsein* is significant in the *Psychology from an Empirical Standpoint*, as it allows a thorny issue in the psychological debate of the time

[3] "I know of examples of young people, desiring to devote themselves to the study of psychology, who, at the threshold of the science, began to doubt their own ability. They had been told that inner observation is the main source of psychological knowledge, and they repeatedly made strenuous attempts at it. But all these efforts were in vain; all they got for their trouble was a swarm of confused ideas and a headache. So they came to the conclusion that they had no capacity for self-observation, which is quite right. But on the basis of the notion which had been imparted to them, they took this to mean that they had no talent for psychological investigation" (42; 22 f.).

to be solved, namely that concerning the presence or absence of unconscious psychic states or acts (143 ff.; 79 ff.). According to Brentano, the hypothesis of an unconscious consciousness, far from constituting a *contradictio in adiecto*,[4] deserved to be examined because it was apparently capable of averting the danger of an infinite regress inherent in the very notion of presentation (*Vorstellung*). If, in an act of presentation, one is conscious both of the content presented (e.g. a sound) and of the act of presenting itself (e.g. hearing a sound), then it seems to be expected that the latter also owes its conscious status to its being presented in a new act of presentation (the presentation of hearing the sound) and so on, according to a potentially unlimited presentations *of* presentations. Precisely in order to prevent such a regression, it seems reasonable to interrupt the series by resorting to an unconscious presentation (171; 94) underlying the primary level of presentation (the hearing).

However, Brentano argues, since "there are undoubtedly occasions when we are conscious of a mental phenomenon while it is present in us," (176; 97) it is perhaps worth taking a different route to escape such a counterintuitive hypothesis. In the wake of Aristotle,[5] Brentano's loophole finds unwarranted such a proliferation of acts, because it does not take into account the mutual inherence between presentation and presented object.

> Rather, inner experience seems to prove undeniably that the presentation of the sound is connected with the presentation of the presentation of the sound in such a peculiarly intimate way that, by existing, it contributes at the same time inwardly to the being of the other. This suggests that there is a peculiar interweaving [*Verwebung*] between the object of inner presentation and the presentation itself, and that both belong to one and the same mental act (179; 98, transl. mod.).

The retrieval of the Aristotelian insight is very fruitful here. It makes it possible not only to sidestep the two horns of the problem – *aut* unconscious

[4] "We use the term 'unconscious' in two ways. First, in an active sense, speaking of a person who is not conscious of a thing; secondly, in a passive sense, speaking of a thing of which we are not conscious. In the first sense, the expression 'unconscious consciousness' would be a contradiction, but not in the second. It is in the latter sense that the term 'unconscious' is used here" (143 n.; 79 n.).

[5] Cf. *De Anima*, III 2, 425 b 12-15: "Since we perceive that we are seeing and hearing, it is necessary that one perceives that one sees either by sight or by some other sense. But then the same sense will be of sight and of the underlying colour, with the result that either there will be two senses of the same thing or a sense will be of itself" (Aristotle 2016: 52; cf. Brentano 1973: 179 n.; Eng. tr. 98 n.). As is well known (cf. *De Anima*, III 2, 425 b 15-17), Aristotle's examination resolves the alternative by ascribing some kind of self-awareness to the percipient sense, precisely in order to avert the regress to infinity: "Further, if the sense which perceived sight were to be other than sight, then either this will carry on into infinity or there will be some sense which will be of itself" (Aristotle 2016: 52; cf. Brentano 1973: 183; Eng. tr. 100).

consciousness, *aut* regress on the side of presentation – but also to root out the presupposition of their bifurcation, identifying the very source of the dilemma. The latter is, in fact, the result of a kind of self-deception or perspective blunder triggered by our analytical commitment. This is how Brentano puts it:

> The presentation of the sound and the presentation of the presentation of the sound form a single mental phenomenon; it is only by considering it in its relation to two different objects, one of which is a physical phenomenon and the other a mental phenomenon, that we divide it conceptually into two presentations (179; 98).

If we succeed in overcoming – and emancipating ourselves from – the analytical import of our reflection on the psychic phenomenon, it is possible to grasp its "dual peculiarity," which consists in having "the sound as content within it and [...] itself as content at the same time." This also permits to consider, within the same unity, the content presented as its *primary object* and the presentation itself as the *secondary object* of the act. As Brentano notes, the primacy of the presented content over the presenting act is not a matter of temporal order. The two objects would be given at the same time as the act is performed. Rather, the primacy relates to the "nature of the case," since if the occurrence of the presented in the absence of the presenting appears "inconceivable," a presenting without the presented would result in "obvious contradiction." Each presentation thus seems to be directed "in the most proper sense of the term" to the presented content, and only "incidentally and as something additional [*nebenbei und als Zugabe*]" to itself (179 f.; 98, transl. mod.).

In the German term *nebenbei*, Brentano translates the Aristotelian expression *en parergo*,[6] emphasizing how the psychic phenomenon is only incidentally or accessorily self-aware, i.e. by reflex or in *a latere* way, with respect to its own intentional theme, namely the object actually presented. It is thanks to this aspect that Brentano is able to distinguish, as mentioned at the beginning, inner consciousness and the perception associated with it from the phenomenon of inner observation or introspection. Although Brentano does not put too much effort to differentiate one from the other,[7] he does proclaim a "universally valid psychological law," that "we can never focus our *attention* upon the object of inner perception" (1973: 41; Eng. tr. 22). He repeatedly rules out a direct observation of the psychic phenomenon, as it would imply a thematization equivalent to that of the experience of a primary object (181; 99).

[6] See how Brentano translates *Metafisica*, XII 9, 1074 b 35-36: "Das Wissen und die Empfindung und die Meinung und das Nachdenken scheinen immer auf etwas anderes zu gehen, auf sich selbst aber nebenbei [ἐν παρέργῳ]" (Brentano 1973: 185; cf. Brentano 1971: 138; Eng. tr. 215).

[7] On the topic see Obstfeld 1983.

Upon closer examination, therefore, in the observation of a psychic phenomenon, we are dealing with a first act directed towards a second act. This can happen through reflection or perhaps through memory, i.e. recalling a previous experience (61; 32). In doing so, however, by deviating from internal perception/consciousness, one ends up shying away from what more than anything else distinguishes it, namely, evidence (128; 70, *passim*). An evidence, that of internal perception/consciousness, which does not need to be justified because immediate, self-evident, and therefore, in its Cartesian inspiration and against all types of skepticism, the ultimate foundation or guarantee of the entire scope of knowledge (198; 109).

All of this leads to a further aspect of the theory of inner consciousness. If we adhere to what was said earlier about the unity of the psychic act, the latter appears as a "characteristic fusion of consciousness and the object of consciousness." Now, to Brentano, this fusion ensures an "inner knowledge" of the act in question, which "contains itself in its entirety as presented and known, in addition to its reference to a primary object" (196; 107, transl. mod.). The shift is significant. Brentano here begins to progressively enrich the prerogatives of inner consciousness: it will no longer be limited to qualifying the psychic act as the mutual inherence of presenting (act) and presented (content) or as the unitary dimension of appearance as such. The purely (self-) manifestative horizon of the inner consciousness will indeed be complicated by the introduction of the functions of judgement (*Urteil*) and feeling (*Gefühl*) into its domain (203; 111). Thus, by virtue of the enrichment of the functions ascribed to inner consciousness, each psychic act can be seen from four different aspects: (1) "as a presentation of its primary object," (2) "as a presentation of itself," (3) as judgmental "self-knowledge" and, lastly, (4) as "self-feeling" (219; 119, transl. mod.).

In this way, the multifarious determination of inner consciousness ends up tracing the three fundamental species of psychic phenomena identified by Brentano in his *Klassifikation der psychischen Phänomene* (1911): presentations, judgements and the class of feelings and will (1971: 125-130; Eng. tr. 206 ff.).[8] When asked about the relationship between the three psychic types of

[8] This conclusion seems to be derived from the approach adopted by Brentano's *Psychology* as early as 1874. By selecting inner perception as the privileged method of investigation, the entirety of psychic phenomena considered by the subsequent taxonomy must somehow already coexist *within* consciousness, at least as a necessity inherent to their possible examination. This aspect does not apply to Husserlian phenomenology, which does not view various types of acts as merely juxtaposed, but rather according to specific relationships of foundation/stratification. Thus, Husserl is not compelled to overload the notion of inner consciousness from the outset with prerogatives of higher cognitive levels such as those related to judgment, thereby preserving its fundamentally manifestative status.

inner consciousness (presentative, judgmental, affective-volitional), Brentano points out that they "are most intimately intertwined," so that "there is a certain ubiquity pertaining to each class in all of our conscious life" (125; 206). On the other hand, he also ascertains "the natural order of the three classes," giving priority to presentation over judgment, and judgment over feeling or will. The presentative function would be the "simplest" and the "most universal," so to speak, "while judgement and love always include a presentation within them. It is likewise the most independent of the three, since it is the foundation for the others." Consequently, it might seem that by virtue of a certain order of foundation one can have purely presentative psychic acts. And yet this is not Brentano's stance, since he does not "want to deny that judgement and love are also represented in some way in every mental state." The relationship among the classes is not so much concerned with the actual concreteness of the *Seelenleben* as with the logical ordering of the more or less complex degree of their thinkability. In other words, it is merely "for a psychological fiction" that "we can conceive, without contradiction, of a being which has no capacity for judgement or love, equipped with nothing but the capacity for presentation, but we cannot conceive of it the other way around" (127 f.; 207).[9]

1.2 *Relations and temporality*

In subsequent works to that of 1874, Brentano gradually refines his conception of inner consciousness, ultimately achieving results that are in some ways paradoxical, given his initial positions. Such developments arise from Brentano's reassessment of intentional reference in terms of his theory of relatives (1971: 133 ff.; Eng. tr. 211 ff.; 1976: 124-138; Eng. tr. 74-82).[10] Given certain similarities that I will later discuss, the psychic reference differs from a relation in the proper sense by virtue of the different existential implication it entails. In the comparative relation "Caius is taller than Titus," the existence of Caius (the *foundation* of the relation) implies the existence of Titus (the *term* of the relation) as a necessary element for the comparison to make sense. This is because, without Titus, the comparison would not be meaningful since the term of comparison would be lost – the relation can be expressed as follows to em-

[9] In an addendum to his *Klassifikation*, Brentano will reconsider the attribution of an "emotional tone" to the inner consciousness of every experience. According to Brentano, there could be visual or auditory experiences detached from volition or feeling (139; 215). However, what remains more problematic from a Husserlian perspective is the persistence of a predicative and thus self-knowing element at such a basic stage as that of inner consciousness.

[10] For a quick overview of Brentano's theory of relations see Kraus (1919: 43-45) and Volpi (1987: 71-73).

phasize the existential implication: "Caius is taller than Titus is" (126; 75). In the case of intentional reference, however, the term of the relation "does not need to exist in reality at all." If "Caius thinks of a centaur," it is rather the existence of the foundation of the relation that is implied by the intentional connection, insofar as thought, in the Cartesian sense, presupposes someone who does the thinking. This allows to distinguish psychic reference from *relativa* in the proper sense, characterizing the former as a "quasi-relational [*Relativliches*]" (1971: 134; Eng. tr. 212).

The commonality between psychic reference and relations is nonetheless reaffirmed by the fact that in both cases, we are dealing (at least) with a dual objectivity instantiated by the two modes of presentation *in recto* and *in obliquo*.[11] From a strictly relational point of view, the foundation is always grasped *in recto*, whereas the term(s) *in obliquo*. The case of psychic reference is trickier because, according to the new 'relational' assessment of intentionality, the act itself is given *in recto* whereas the object is presented *in obliquo* (1971: 145; Eng. tr. 219). This result, however, fails to comply with the previously mentioned distinction between the primary and secondary object of presentation. After all, Brentano departs from the distinction illustrated in the *Psychology* of 1874 between perception and observation, risking to make inner perception a sort of thematic self-observation.

Other passages testify to this change in thought of the late Brentano. In particular, the possibility of directly perceiving oneself is strongly asserted in a text where the modes of presentation *in recto* and *in obliquo* are used to illustrate the perception of time (1974: 37-52; Eng. tr. 28-38). This connection with temporality allows to underscore a final aspect of Brentano's theory of inner consciousness. At some point, he describes "the situation in such a way that we, as sensing beings, are the only things that are sensed *in recto* while our external objects are sensed solely *in obliquo*" (37 f.; 28). It could be argued that Brentano's suggestion in these pages is to adopt the point of view of the psychologist, i.e. the one who perceives from within, by explicitly addressing his own lived experience. "If we carry this interpretation a little further, we will find that the external object of sensation, in addition to having a mode of presentation other than the *modus rectus*, can exhibit various aspects in its modes of presentation", that is, in the present case, its "*temporal differences*" (38; 29).

[11] "At least" because the oblique mode of presentation can be iterated and thus multiplied indefinitely. For example, I can think *in recto* of Caius thinking *in obliquo* of Titus (first-degree) who in turn thinks of Sempronius (second-degree), and so on (1976: 131; Eng. tr. 78). Similarly, considering a comparative relation, one can say "Caius is taller than Titus who is taller than Sempronius, etc." As Brentano elsewhere notes, "the *modus obliquus* itself is not really one single mode; it is differentiated in various ways" (1971: 145; Eng. tr. 219).

According to this perspective, it is possible to understand how *in obliquo* I experience objects as past, present and future on the basis of an implicit reference *in recto* to myself as present: "I never affirm something as having existed yesterday without affirming myself as existing today, and the latter happens *in modo recto*, while the former *in modo obliquo*" (41; 31). The *modo recto* of presenting myself as present becomes, therefore, the foundation of the temporal determinations presented *in obliquo* as terms of the intentional relation.[12] We are thus witnessing the reversal of the thesis of 1874 regarding the relationship between the primary and secondary object of the act:

> the proof that inner perception is limited to the present and does not relate in the slightest to the past or the future […] makes it immediately clear that mental phenomena, like physical phenomena, can be presented and judged not just in a secondary but in a primary manner as well (1974: 15; Eng. tr. 12).

This line of reasoning implies a significant consequence as far as inner consciousness is concerned. Being called upon to play the role of foundation *in recto* for the grasping of temporal differences *in obliquo*, the inner consciousness has to be limited to the punctual dimension of an instantaneous present: "our inner perception, which perceives with evidence, is entirely restricted to a single now and presents us with certainty only with one boundary which is in the present" (1976: 130; Eng. tr. 77). In order to fully understand this point, we must recover the notion of original association (*ursprüngliche Assoziation*) or *proteraesthesis*, the real keystone of Brentano's account of temporality.[13] Our intuition of time depends on the fact that the perception (*aesthesis*) of something present (e.g. a sensation) is immediately associated with its modified perception as no longer present. This original modification or association – a fundamental psychological law – takes the name of *proteraesthesis*, which literally means: perception, sensation (αἴσθησις) of what comes before (πρότερος). To the real content perceived as present is thus associated its *derealizing* aspect,[14] namely its being apprehended now as what was previously present. In this way, thanks to the

[12] See Brentano (1976: 131; Eng. tr. 78), where declinations of time such as "before" and "after" are illustrated by analogy with the possibility of iterating the terms of a relation at will.

[13] The reflection on temporality encompasses the entire span of Brentano's production. Brentano developed an interest in temporality from the outset through a close examination of Aristotle (Fugali 2005: 100-115; de Warren 2018). On this basis, he then embarked on an autonomous path encountering numerous stages of elaboration and reconsideration. Volpi (1987: 77 ff.) reconstructs at least four phases, whereas Fugali (2005: 214-215 n.) even enumerates six of them. Regarding the Brentano-Husserl relationship on the question of temporality, see also Kraus 1930, Albertazzi 1990/1991; Rinofner-Kreidl 1995/96; Benoist 2008; de Warren 2009: 97-140.

[14] The derealizing character of *proteraesthesis* is justified in Brentano's view as a modification perpetrated by the imagination.

association between *aesthesis* and *proteraesthesis*, the presentation of the object acquires a temporal depth due to the now stretching into the just past.

Relating what has just been said to the theme of our investigation, we can see that the phenomenon of *proteresthesis* only concerns external perception and not inner consciousness (Brentano 1974: 50 f.; Eng. tr. 37 f.). Inner consciousness does not have any *proteraesthesic* character, any temporal depth except that of a liminal present serving as reference point for the temporal determinations of external perception. Without such a point of reference "it would then be impossible to experience outer phenomena as past" (Brentano 1971: 224; Eng. tr. 257). Besides, to attribute a temporal modification to the inner side of the psychic act would lead us back to an infinite regress.

> The evidence of inner perception is therefore, wholly limited to the present. This is true for inner sensation in general. It tells me only that I now have this aesthesis together with a continuous proteraesthesis of physical phenomena. If it were itself accompanied by inner proteraesthesis as well, we could easily prove that this would lead to infinite complications, for its object would also have a proteraesthesis and so would the proteraesthesis, *ad infinitum*; this would result in a kind of continuum of an infinite number of dimensions (51; 38).

If, as Blumenberg once said, "the phenomenologist is defined by fearlessness in the face of infinite regress" (1986: 374, my transl.), Brentano seems to abstain and recoil from such a possibility, precluding himself from exploring a path full of consequences. Husserl's way out, *viceversa*, will consist precisely in anchoring inner consciousness to that mechanism of intentional modification of the now, bearing the name of retention.[15] Through the retentional modification, he will accomplish, not by chance, a multidimensional complication of the temporal *continuum* of the stream of consciousness (1966: Nr. 32). The risk of regress will constantly loom like a specter in the most enigmatic depths of Husserl's speculation, repeatedly dazzling every attempt of a solution (1966: 29, 114; Eng. tr. 31, 119).

The aim of this brief and selective examination of Bretano's writings was not to merely reconstruct the theoretical background from which Husserl inherits the notion of inner consciousness. The main intent was indeed to outline the horizon of philosophical dilemmas that the retrieval of such a notion implies. Whether one resorts to the hypothesis of an inner consciousness as an alternative to the existence of unconscious psychic states, whether in attempting

[15] This will lead to Husserl's theory of the double intentionality of retention (*Quer- und Längsintentionalität*) and to assimilate inner consciousness to the question of the self-constitution of the stream of consciousness (Husserl 1966: § 39, Nr. 45).

to philosophically develop this hypothesis one runs the risk of objectifying lived experience or facing the danger of infinite regress, whether inner consciousness seems to entail a wholly peculiar mode of intentional reference (*en parergo*) inoculating the presumed guarantee of absolute evidence, and finally, whether inner consciousness proves to be inherently related to the similarly relational theme of temporality – all of this would have to be reconsidered and debated within the new perspective opened up by Husserl's phenomenology. The nature of the issue is such that, while taking different paths, the problem seems to break down into the same fundamental stages.

2. *Husserl's reception in the* Logical Investigations

2.1 *Inner and outer perception in the* Beilage

At the time of the first edition of the *Logische Untersuchungen* (1900-1901), Husserl's position seems to oscillate unresolvedly between intentions of detachment and unconscious adherence to Brentano's paradigm. The fierce critique of psychologism conducted in the *Prolegomena* does not appear to have been fully integrated into the subsequent *Investigations*. Despite the initial declaration in favor of a complete "freedom from presuppositions" (1984: 24; Eng. tr. 177), at the turn of the century Husserl's approach is still tainted by some residual psychological commitments, tied to a conception of the *Erlebnis* understood as a psychic, i.e. empirical, fact.[16]

With regard to the topic at hand, as early as the Introduction to the *Investigations*, Husserl complains of the "false antithesis between inner and outer perception" (1984: 13; Eng. tr. 170, transl. mod.) endorsed by psychologists. As emerging from the *Beilage* specifically devoted to this opposition, the trouble

[16] These aspects can be appreciated by scrutinizing the occurrences, in the first edition of the *Investigations*, of the word *real*, which refers to the factual existence of the *Erlebnisse* and the empirical status of the Ego, in contrast to *reell*, as amended in the second edition, which denotes what is purely effective in the lived experience (e.g. its content) (1984: 15, 57, 129, 357, 362-364). Moreover, in the first edition, the adjective *rein* does not refer to the sphere of pure consciousness as attainable through the phenomenological reduction, but to the strictly *descriptive* psychological analysis, as opposed to the *genetic-physiological* one (cf. 357, 411 n. for some relevant additions to the second edition). The first clear separation of phenomenology from descriptive psychology is attested in 1903, in a review of an article by Th. Elsenhans (Husserl 1979: 206). Besides, as Husserl would say in 1913, it is only around 1907/08, after the refinement of the method of reduction, that the overcoming of psychologism can be said to have been fully achieved (2002a: 324 and n. 4; Eng. tr. 59). For all this, I rely on Lavigne's (2005: 118-129) meticulous examination. As the author aptly notes (143), any attempt to conceal such psychological commitments prevents an understanding of the need for change and development in Husserl's thought in the early years of the century.

lies in the definition of the psychic as a phenomenon proper to inner perception and the consequent identification of the latter with inner consciousness (756 n.; 358 n. 2).[17] Even before Brentano, the critical excursus of the *Beilage* directs its arrows both at the history of modern thought (Descartes, Locke) – guilty of becoming entrenched in dichotomies such as *mens* vs. *corpus* or sensation vs. reflection – and at the naive standpoint of the common sense, in which these oppositions echo in the form of distinctions such as those between *external perception* and *perception of the self*, *sensuous* and *internal perception* (1984: 751 f.; Eng. tr. 335 f.). Husserl's aim, then, is akin to someone preparing to fluidify these oppositions, noting how they intersect with other, more fundamental ones, such as the distinction between perception endowed with *adequate* or *inadequate evidence*, or that, as yet embryonic, between *immanence* and *transcendence*.

Focusing on Brentano, here as elsewhere, Husserl reproaches the master for failing to distinguish between the content and the object of consciousness.[18] Consequently, he argues, Brentano did not realize that the evidence of inner perception can also be attributed to external perception, provided that it is ascribed to the "lived sensuous content" and not to the object. In this respect, Husserl emphasizes that "there are, therefore, evident perceptions of 'physical' contents, as well as of 'psychical'" (1984: 768; Eng. tr. 345, transl. mod.). After all, it is the division between inner and outer perception, carried out by Brentano along the lines of the distinction between psychic and physical phenomena, that must be called into question. To Husserl, in fact, the experience of psychic states is not necessarily evident, "since these are perceived with a bodily location." The state of pain inflicted by a dental abscess is a psychic phenomenon that contains an implicit reference to a transcendence, even if only of a bodily nature (760 f.; 340 f.). Such transcendence could therefore compromise the evidence of a perception internally directed to the pain suffered. Thus, Husserl concludes, "it is absolutely clear that the conceptual pairs of inner and outer, and of evident and non-evident perception, need not coincide at all" (769; 345).

As well known, Husserl's criticism resolves the false opposition between inner and outer perception, based on the fracture between the psychic and the physical, through the recovery of a more genuine "*gnoseological* difference [erkenntnistheoretischen *Differenz*]" between perceptual modalities endowed with adequate or inadequate evidence. In adequate perception content and object coincide in the immanent fullness of the lived experience, whereas inadequate perception grasps the object as presumptive insofar as the latter tran-

[17] Regarding the *Beilage* see de Warren 2003.

[18] Distinction that Husserl had instead found in Twardowski 1894.

scends the content through which it is apprehended (769 f.; 346, transl. mod.). The argument in the *Beilage* would be far more detailed, but for the purposes of this paper, it is sufficient to suggest the following on the basis of the above. The critique of Brentano's distinction between two perceptual directions (inner and outer) does not lead Husserl to the rejection of the notion of inner consciousness, but to a different way of conceiving it. For the time being, it is crucial to consider that "he [Brentano] has employed an essentially distorted concept of perception in the case of inner perception" (769; 345).

2.2 *Inner consciousness and lived experience in the* Fifth Investigation

The *Fifth Investigation* opens with the distinction of three different concepts of consciousness:

> 1. Consciousness as the entire, phenomenological being of the spiritual Ego. (Consciousness = the *phenomenological Ego* as the 'bundle' or the interweaving of psychic experiences).[19] 2. Consciousness as the inner awareness [*als inneres Gewahrwerden*] of one's own psychic experiences. 3. Consciousness as a comprehensive designation for 'mental acts', or 'intentional experiences [*Erlebnissen*]', of all sorts (1984: 356; Eng. tr. 81).

While the first concept denotes the whole life of the Ego, which in the second edition would be referred to as the *Einheit des Erlebnisstroms*, the second and third represent respectively the inner consciousness and the acts or lived experiences that make up the stream itself. The distinction between these two last meanings can be clarified as the distinction between an *intransitive* and a *transitive* sense of being conscious. Inner consciousness would thus be simply (self-)conscious (rather than unconscious), whereas the lived experience would be conscious of something else.[20]

The first concept reflects, we might say, the general project of the time (of clear Brentanian and, in this case, anti-Aristotelian inspiration) of a "psychology without a soul [*Psychologie ohne Seele*]" (Husserl 1984: 370 f.; cf. Brentano 1973: 16; Eng. tr. 8).[21] This project would consist of an analysis of the succession of psychic phenomena inspired by the empiricist image of the bundle (*Bündel*), free from any (neo-)Kantian assumption concerning an alleged egological principle (pure ego) ensuring the unity of the flow.[22] In the wake of

[19] I have modified the English translation of the first concept, in accordance with the first edition of the *Investigations*.

[20] Cf. Zahavi 2002: 52. More generally, on the *Fifth Investigation* see Cobb-Stevens 2003.

[21] As stated by Brentano, the expression was coined by Lange 1866: 465.

[22] Regarding the "psychology without soul," see Husserl's new position (1976: 195 f.; Eng. tr. 206 f.). Also, refer to §§ 4 and 8 of the *Fifth Investigation*, as well as the respective notes, for the polemic

Natorp's *Einleitung in die Psychologie nach Kritischer Methode* (1888), Husserl was also already aware of the distorting effect of reflecting on the psychic act.

> We perform the description after an objectifying act of reflection, in which reflection on the ego is combined with reflection on the experienced act to yield a relational act, in which the ego appears as itself related to its act's object through its act. Plainly an essential descriptive change has occurred. The original act is no longer simply there, we no longer live in it, but *we attend to it and pass judgement on it.* We must therefore avoid the misunderstanding which our present discussion has just ruled out, that of treating relation to an ego as of the essence of an intentional experience itself (1984: 391; Eng. tr. 101).

It would be this reflective transfiguration of the *Erlebnis* that triggers the illusion of a reference to the act (thus objectified) by an egological pole. Thus, this reference does not appear to be an essential trait of lived experiences *per se*, but rather a characteristic that overdetermines our conception of the act, once the latter is reflectively placed within an intentional, already polarized context, according to the classical subject-object opposition.

Similar observations are found in the discussion of the other two meanings of consciousness outlined above. A general tendency of the *Fifth Investigation* is indeed a revision of the inherited psychologistic lexicon, aiming at a new terminological refinement. As stated at the beginning of § 11, "while we adhere to Brentano's essential characterization [of psychical phenomena], our departures from his opinions force us to abandon his terminology" (384; 97).

Terms like (re)presentation (*Vorstellung*), phenomenon (*Phänomen*), lived experience (*Erlebnis*), perception (*Wahrnehmung*), manifestation (*Erscheinung*), if uncritically assumed can lead to misunderstandings (*Mißdeutungen*). We could say that the *trait d'union* of these criticisms lies in the constant inability to distinguish, within the phenomenal domain, a sense whereby something *is given* or *appears* (erscheinen) *as an object*, and one whereby a consciousness or a lived experience *lives itself* (sich erleben) *as a subject*. A phenomenological critique of knowledge should in fact clarify how: "as belonging in a conscious connection, we live the appearing of things, as belonging in the phenomenal world, things appear before us. The appearing of the things does not itself appear to us, it is lived through [*erlebt*]" (359 f.; 83, transl. mod.).

With regard to the different meanings of the notion of appearance (*Erscheinung*), it is significant that for Husserl it denotes both the object that manifests itself and the *Erlebnis* through which it appears, *although not in the same sense*;

with Natorp concerning the pure ego and the reassessment occurred in the second edition of the *Investigations*.

finally, the *Erscheinung* does not concern, except in a misleading way, "the *real* [reellen] *constituents* of appearances" such as sensations (762 f.; 341 f.). As a result, in § 5, Husserl can charge Brentano the conflation of two senses of consciousness (as lived experience and as inner consciousness), equating them as if they were two different directions of the same act of perception (a primary one, directed towards an external object, and a secondary one, directed towards an internal object). As Husserl surmises, Brentano's misconception could be traced back to "the equivocation which pushes us to treat consciousness as a sort of knowing, and in fact of intuitive knowing" (366; 476).

Such knowing, if I may further elaborate, would be the one stemming from an immediately cognitive and therefore not simply manifestative conception of the *Erlebnis*, a conception conveyed more by Brentano's lexical choices than by his real intents. I have previously outlined Brentano's original intention to trace a distinction between the purely presentative character of *innere Wahrnehmung* and the objectifying character of *Beobachtung*. Nevertheless, we have seen how, in later works, the inadequacy of this purpose has led to a turnaround, namely the possibility of perceiving the *Erlebnis* no longer *en parergo* but *in recto*. One could therefore venture, given the importance of lexicon, that such a conclusion was already inherent in Brentano's terminological choice to qualify the presented and the presenting as two objects, one external and one internal, of perception.

Beyond the lexicon, however, it is Brentano's stance, which conceives inner consciousness as *cognitively mediated* by an act of perception (albeit understood only as an internal inflection of lived experience and not as an additional act), that proves to be questionable for Husserl. In fact, what the remarks in the *Fifth Investigation* and the *Beilage* tend toward is a preliminary characterization of the immanent, subjective dimension of experience.[23] Husserl begins here to discern one of the fundamental structures of phenomenal experience: phenomena are not only intended as bearing their own sense (*as* an object, *as* something), but they also presuppose the reference to a consciousness for which this sense acquires a possible meaning. Phenomenal appearance is thus always an appearing *to/for* someone, but (self-)experiencing this someone cannot be the same as experiencing something as an object. The act is innerly and immediately *lived through*, it is conscious (in the intransitive sense of the term) without the need for any perception directed towards it that reifies it as an *immanent object*.[24] "It is in this sense – Husserl says – that *what* the ego or consciousness lives, is its lived experience" (362; 85, transl. mod.). As stated in the *Sixth Investigation*: "To be lived is not to be objective [*Erlebtsein ist nicht*

[23] Cf. de Warren 2003: 156; Zahavi 2002: 58 f.

[24] *Immanente Objeckt* is another misleading expression criticized in § 11 of the *Fifth Investigation*.

Gegenständlichsein]" (669; 279, transl. mod.).

Now, what has been noted allows for a correct assessment of a couple of passages in which – at least apparently – Husserl seems to throw the baby out with the bathwater, that is, he seems to dismiss *tout court* the idea of inner consciousness, and not merely reject Brentano's terminology and the conception it conveys. In the first passage, it is stated that "the need to assume the continuous activity of inner perception cannot be phenomenologically demonstrated" (367; 87, transl. mod.). In the second, Brentano's view is reported only to be openly denied: "'every psychical phenomenon is an object of inner consciousness.' We have already mentioned the grave misgivings that keep us from assenting to this" (384; 98). Both passages aim not so much at the rejection of the notion of inner consciousness – its relevance will always be upheld in subsequent works – but at the rejection of a certain way of grasping and conceptualizing it, i.e. *as perception* and *of an object*.

In fact, Husserl remolds here expressions such as 'inner perception' and 'immanent object.' They end up more appropriately denoting the kind of *reflective* thematization of the *Erlebnis*, abstracted from the concreteness of the stream and objectified, that Brentano had reserved for the *Beobachtung*: "'reflection', however, implies that what we reflect upon, the phenomenological lived experience, is rendered objective to us (is inwardly perceived by us)" (669; 279, transl. mod.). Husserl's goal is not to banish reflection and the perceptual thematization of lived experiences from the realm of phenomenology. On the contrary, reflectivity will turn out to be a core element of the method and before that a constituent of objectivity itself (14-16; 170 f.).[25]

On closer examination, Husserl's aim is twofold. On one hand, he seeks to safeguard the *Gewahrwerden* of inner consciousness, stripping away Brentano's heritage of all its terminological ambiguities (internal perception, immanent object, etc.); in so doing, he maintains the genuine meaning of the *Erlebnis* as a form of original living, prior to all derived modes of perceiving, attending to, noticing, and reflecting upon lived experiences. On the other hand, these derived modes of intentionality are not banned from phenomenological research – as Brentano instead claimed to dismiss observation from descriptive psychology – but are rather included as steps of a graduated path of objectification. Similarly graduated should be understood the subjective aspect of this path[26] – hence Husserl's ambivalence towards the ego, now deemed superfluous, now necessary, depending on the level, the context of analysis

[25] Cf. also Husserl 1976: § 77.

[26] In 1908 Husserl (2003: 70) will eventually and properly distinguish the two interrelated processes of *Objecktivierung* and *Subjecktivierung*.

at hand. Pointing out, however, that the subjectification of the ego is clearly distinguished from the objectification of thingly (*dinglich*) reality, Husserl concludes:

> As well as the turning of noticing to a thought, to a sensation, to a feeling of unease, etc., makes these lived experiences objects of inner perception, without making them objects in the sense of things, so well would that centre of relationship, the ego, and every particular relationship of the ego to a content, as noticed, also be objectively given (1984: 373; Eng. tr. 92, transl. mod.).

2.3 *Beyond the* Logical Investigations *(1902-1904)*

In the years immediately following the *Investigations*, Husserl's commitment to a renewed "inner gnoseological work" (2001: 10) seems to confirm the path taken in his "break-through" book (1975: 8; Eng. tr. 3): the advancing erosion of the paradigm of descriptive psychology, despite the seemingly unchanged adherence to it. The course held at Göttingen and titled *Allgemeine Erkenntnistheorie* (*WS* 1902-03) provides a striking example of this trend. In these lectures, Husserl undertakes a first development of his own method which will culminate in the overcoming of the idea of phenomenology as descriptive psychology. Although the latter still represents the foundation of the theory of knowledge (2001: 69) we see a redefinition of its prerogatives.

First of all, Husserl points out the merely empirical character of psychology, which only deals with "facts of consciousness," drawing them from individual experiences and generalizing inductively (27). On this basis, a proper phenomenological method aimed at the clarification (*Aufklärung*) of cognitive operations through a renewed understanding of the *Erlebnis* stands out (60). The latter will no longer be understood as a psychic-empirical fact, but elevated to a "single case [*Einzelfall*]" of a true "essence analysis" (66, 77).[27] For the first time, the descriptive tenor of psychology undergoes here a rudimentary purification. The *Erlebnis* is examined as such, abstracted "from all empirical objectifications" and presuppositions, that is from everything that transcends its real (*reell*) content, including the reference to the empirical ego or the objective time[28] still found in

[27] All excerpts from Husserl's untranslated works are my translations.

[28] The question of temporality is telling about the development of Husserl's thought. Whereas at the time of the *Investigations* (1900-01) the analysis is carried out "under the presupposition of the objective time in which psychic experiences flow" (1966: 160; Eng. tr. 164), in 1904 Husserl's point would be "the complete exclusion [*Ausschluß*] of all suppositions with respect to objective time" (187; 193). As we shall see, this change stems from the purification of phenomenological analysis due to a first form of reduction, which resolutely purges all the empirical-factual presuppositions that, more or less implicitly and despite their *Voraussetzungslosigkeit*, still tainted the *Investigations*.

the *Investigations*. A "purely immanent description" is now directed to the lived experience, which is raised "as generality," as "lived-experience-idea [*Erlebnis-Idee*] or lived-experience-essence [*Erlebnis-Wesen*]" (2001: 77 f.).[29]

The purification of the empirical character of the *Erlebnis* is due to a form of eidetic reduction already at work in the *WS* 1902-03. The term *Reduktion* makes its appearance here in two distinct occurrences belonging to the same page. First, when speaking of the "ideal possibility of an adequate fulfillment" of the object, Husserl writes: "this reduction to the ideal possibility is the foundation for our theory of knowledge" (2001: 198 f.). A few lines below, however, we find the meaning attributed here to the reduction clarified:

> It is not about the person or their current act – there is no question of it at all, it is excluded [*ausgeschlossen*] by the phenomenological reduction – but about the general essence of this presentation, of its general species, and to this belongs the idea of fulfillment as a general possibility independent of all empirical things (199).

Although Husserl qualifies it as "phenomenological," what one finds here is a reduction that, following Lavigne (2005: 295 ff.), can be defined as "restrictive." In fact, it does not open the view to phenomenality as such, as will be the case with the proper phenomenological reduction from 1906-07 onwards. On the contrary, it rather significantly reduces it, by narrowing the field of inquiry to only those components that are truly immanent (*reell*) to lived experience.[30] Besides, a specific task is thus fulfilled. Emancipating Husserl's analysis from all the empirical implications (the factual nature of the *Erlebnis*, the reference to the transcendent object, to the empirical ego, to objective time, etc.), this kind of reduction distances phenomenology from descriptive psychology and opens the door to the "eidetic"

[29] As observed by McDonnell (2011: 103, 108), it is precisely this eidetic tenor that constitutes the scientificity of the phenomenological endeavor as opposed to the psychological-descriptive one, which relies on the conveyed evidence of inner perception. Since the latter, to Brentano, is a non-attentive and merely incidental type of self-awareness, it is unclear what methodological contribution it could represent for a scientific enterprise of any kind. However, as McDonnell (99) suggests, Husserl's landing on transcendental phenomenology will also be due to a radicalization or a fully "transvaluation" (*Umwertung*, cf. Husserl 2002a: 272; Eng. tr. 16) of inner consciousness conducted in his time analyses (see my Nobili 2022: §§ 1.3, 2.2 for these developments).

[30] At the time, Derrida (1990: 148) had already emphasized the "privative" character of the first kind of reduction practiced by Husserl. Costa (1994) and Lohmar (2012) suggest that this type of reduction is already at work somewhere in the *Logical Investigations*. However, this interpretation seems to be supported only by the retrospective account that Husserl provides with the revision for the second edition of the *Investigations* (Husserl 1984: 368 f., 413 n.; Eng. tr. 88, 354 n. 26). In the first edition, the occurrences of the verb *reduzieren* (370, 381, 395) in no way denote a technical use of the term. Even where it might appear to be the case, namely where Husserl speaks of a "*phänomenologisch reduzierte Ich*" (364 n.), what he has in mind is the distinction between the psychical reality (*Realität*) of the ego and its bodily component, and the emphasis on the former as a specifically phenomenological (i.e., for the time, psycho-descriptive) field of research.

and "immanentist" researches of the years immediately following (1904-1905).[31]

With regard to the main topic of *innere Bewußtsein*, the course on the *Allgemeine Erkenntnistheorie* does not make any significant changes compared to those found in the *Investigations*. The purpose of a "return to inner intuition" (2001: 60) is reaffirmed as the foundation of phenomenological analysis, as well as the cogitative evidence of such intuition and the capacity of inner perception to reflect on one's own lived experience. Yet, keep in mind, that "in this, we are not pursuing thingly knowledge; we are also not exploring the soul as a supposed thing to which lived experiences adhere as accidents" (80). Rather, the reflective import of inner perception must be understood as a "mere acknowledgment of what is actually given" (94).

Husserl will similarly expound two years later in the first part of the course *Hauptstücke aus der Phänomenologie und Theorie der Erkenntnis* (*WS* 1904-05). In §§ 5-6 he carries on the revision of the psychological lexicon inaugurated in the *Beilage* (to which he expressly refers: cf. 2004: 22). The methodological development initiated in the meantime allows Husserl a greater awareness both conceptually and expressively. The eidetic-immanentist turn is fully affirmed here insofar as "we abstract from the transcendent intentions, 'we exclude them'" (20 n.). The immanentist restriction of phenomenology to the real (*reell*) component of lived experience, to the content of consciousness, reveals a coincidence between the presenting and the presented that is lacking in the case of external perception (19). It is only this coincidence and not the nature of the object in question that ensures the (adequate) evidence of *Wahrnehmung*. The restriction of phenomenology to the lived content is what in fact enables inner "perception," dictating its peculiar modes of exercise. Its reflective tenor must indeed be distinguished both from the perception of a thing-like nature and from that attributed to the inner sensory experience within the act.

[31] As one of the anonymous reviewers of this paper (to whom I am grateful) pointed out, it is probably true that some of the analyses in the *Investigations* may aspire – *before* Husserl's revision – to this eidetic and immanentist status. However, I am persuaded that *before* the explicit theorizing of the reduction some residual psychologistic commitments still corrupt the purity and the truly immanent status of the phenomenological *Betrachtung*. See however Husserl's ambivalent self-criticism in 1913: "*De facto* the analyses [in *LI*] were carried out as *analyses of essences* but not everywhere in an equally clear, reflective consciousness. The entire refutation of psychologism is based on the fact that the analyses (especially those of the *Sixth Investigation* but also the others) are claimed to be analyses of essences, hence apodictically evident analyses of ideas. In general, however, I did not want to concede to myself that what I for many years had looked upon as far as psychology was concerned as derived from inner 'adequate perception' should now all be a priori, or, be comprehensible as such" (2002a: 312; Eng. tr. 51).

If, however, in phenomenological analysis, we look purely at sensations, then they are valid for themselves. A new perceptive direction is constituted, which aims [reflectively] at what is lived as it is given. It is indeed important to understand clearly that sensations, in external perceptions, are lived but not perceived, and that moreover, when we direct our gaze to them, this takes place in new [reflective] perceptions that have a totally different character from that of the original perceptions, the external perceptions (19 f.).

The restricted account of the act redirects reflectively the gaze towards the immanent components of the *Erlebnis*. Thereby, the unique dimension of sensing (*Empfinden*) becomes manifest. Sensing cannot be regarded as an act of perception (whether external or reflective) that intends something immanent in a thematic, attentive manner: "we do not turn our gaze 'intending [*meinend*]', 'attentive' to the contents of sensation." This new dimension is not that of an intentionally performed act: "sensing [*Empfinden*] <is> not a giving 'act'" and, indeed, precisely "on it the apprehension is built, the consciousness of transcendence, in which the 'intending' lives." The sensory content is rather that upon which apprehension exercises its animating function; it is what substantiates the life of intentionality: "in 'adequate perception', intending is addressed to the contents of sensation, it lives into the sensing [*es lebt sich ein in das Empfinden*]" (20 n.).

In light of the above, it should be noted that the use of the term "perception," even in quotation marks, is entirely equivocal here. We have seen that the dimension proper to sensing, given the indistinction in force between the presenting and the presented, does not conform to the structure of an act of perception. However, what truly matters here is the qualification of the term perception – the meaning of which varies in correlation to the type of object perceived.[32] It is the contextual specification of the term that allows to finally distinguish: (1) the adequate "perception" of sensory contents (the sensing) from (2) the inner "perception" of the act (the reflection), and thus to discern both from (3) the external perception of a transcendent thing (perception commonly understood).

Let us recap the movement of thought that led Husserl to this outcome. The necessity to free phenomenology from the conceptual equivocations conveyed by Brentano's terminology promotes the epistemological rupture triggered by a form of reduction coinciding with the (reflective) turning of the gaze and the consequent *purification* and *restriction* of the inquiry to the immanence of lived experience. This latter reveals, in turn, the original dimension of the intentional living (the *Erleben*), which, if taken *per se*, is independent of any act

[32] "*Perception* and *object* are concepts that cohere most intimately together, which mutually assign sense to one another, and which widen or narrow this sense conjointly" (Husserl 1984: 666; Eng. tr. 277).

structure or reference to an ego. Moreover, the *Erleben* is independent of any thematic knowledge directed towards it (at the risk of unnecessarily proliferating the number of acts involved):

> If sensing means the intentional living [*das Erleben*], then the content of sensation or sensation [*Empfindung*] means the content itself, in and for itself, without regard to the ego and the nexus in which it stands in the ego. Of course, there is no difference between the act and the object of the act. *The intentional living is not an act*, but expresses the being of the content in the nexus of consciousness. And this being is not *being-conscious as being-known* [Bewusstsein als Gewusstsein] (infinite entanglement of act [*unendliche Aktverwicklung*]) (Husserl 2004: 23 f.).

Here, one should appreciate how the eidetic-immanentist refinement of the analysis – the reduction to the purely immanent (*reell*) components of the *Erlebnis* – allows Husserl to solve the critical issues regarding the Brentanian legacy raised since the *Investigations*. Radicalizing Brentano's conception that still conceived inner consciousness according to the subject-object opposition, Husserl disclosures a more original dimension of intentional life.[33] Ultimately, the *Erleben* is characterized by (1) the fundamental indistinction between the presented content and the presenting lived experience, (2) the rejection of any objectifying declination of this coincidence (inner consciousness as a secondary object), and finally, (3) the priority over a fully structured intentional act. Indeed, Husserl is here delineating a level of consciousness preordained with respect to that of the act, a level that is not already structured according to the polarity of a dual reference: to the object and to the ego. The *Erleben* then corresponds to a pre-intentional dimension, upon which it becomes conceivable the unfolding of the properly called act-intentionality (*Akt-intentionalität*). The thinkability of such articulation will arise, from the standpoint of Husserl's phenomenology, by penetrating the realm of inner time-consciousness with its stream-"intentionality" (*Strom-"Intentionalität"*).[34] But here another story begins and our prehistory ends.

[33] If, on one hand, Brentano's conception of *innere Bewußtsein* avoids the harmful consequences of a reflective model of self-consciousness – specifically, avoiding the Scylla of an act reflecting on another act and thereby conferring upon it the character of self-consciousness (and triggering a reflective regress of acts directed towards other acts) – on the other hand, this conception falls into this very model by failing to avoid the Charybdis of a conception based on the subject-object opposition (cf. Zahavi 1998). In contrast to this, Husserl's proposal manages to avoid both aspects of the reflective model, thereby dodging the critiques levelled by the opponents of such a theory of self-consciousness (D. Henrich, K. Cramer, M. Frank, U. Pothast); see the classic study by Zahavi 1999.

[34] For this distinction cf. Husserl 2002b: 183.

Conclusions

As we have seen, the way Husserl ends up transvaluating the notion of *innere Bewußtsein* is part of his critical appraisal of the inherited psychologistic lexicon. Since the first edition of the *Logical Investigations*, his endeavor tends to overcome Brentano's view, in which inner consciousness directly relies on the inner perception of mental states. By virtue of this reappraisal, a certain confusion among different cognitive objects (primary and secondary) and performances (self-awareness, perception and reflection) is avoided, thus attaining a better evaluation of the intentional life. According to Husserl, inner consciousness has to be sharply severed from inner "perception." Whereas the latter turns out to represents the reflective thematization of the *Erlebnis* – attributed by Brentano to inner observation and banned from the realm of descriptive psychology – the former denotes the true inner life of consciousness. This is what Husserl calls the *Erleben*, i.e. that original dimension of sensing or living through its own contents, which is logically prior to the actual, properly intentional perceiving of any objects. Stemming from an early improvement of Husserl's own method, able to split the project of phenomenology from that of descriptive psychology, this outcome would be crucial for his later time-analyses.

Filippo Nobili
University of Palermo
filippo.nobili@unipa.it

References

Albertazzi, Liliana, 1990/1991, "Brentano, Meinong and Husserl on internal time", in *Brentano Studien*, 3: 89-109.

Aristotle, 2016, *De Anima*, Eng. tr. by Christopher Shields, Clarendon, Oxford (UK).

Benoist, Jocelyn, 2008 "Modes temporales de la conscience et réalité du temps. Husserl et Brentano sur le temps", in Benoist, Jocelyn, ed., *La conscience du temps*, J. Vrin, Paris: 11-28.

Blumenberg, Hans, 1986, *Lebenszeit und Weltzeit*, Suhrkamp Verlag, Frankfurt am Main.

Brentano, Franz, 1971, *Psychologie vom empirischen Standpunkt.* Zweiter Band. *Von der Klassifikation der psychischen Phänomene*, Felix Meiner Verlag GmbH, Hamburg; Eng. tr. by Antos C. Rancurello, Dailey B. Terrell, Linda L. McAlister 1995, *Psychology from an Empirical Standpoint*, Routledge, London-New York.

—, 1973, *Psychologie vom empirischen Standpunkt.* Erster Band, Felix Meiner Verlag GmbH, Hamburg; Eng. tr. by Antos C. Rancurello, D.B. Terrell, Linda L. McAlister 1995, *Psychology from an Empirical Standpoint*, Routledge, London-New York.

—, 1974, *Psychologie vom empirischen Standpunkt.* Dritter Band. *Vom sinnlichen und*

noetischen Bewußtsein: Äußere und innere Wahrnehmung, Begriffe, Felix Meiner Verlag GmbH, Hamburg; Eng. tr. by Margarete Schattle, Linda L. McAlister 1981, *Sensory and noetic consciousness: psychology from an empirical standpoint III*, Routledge & Kegan Paul, London-Henley.

—, 1976, *Philosophische Untersuchungen zu Raum, Zeit und Kontinuum*, Felix Meiner Verlag GmbH, Hamburg 1976; Eng. tr. by Barry Smith 2010, *Philosophical investigations on space, time and continuum*, Routledge, Abingdon-New York.

—, 1982, *Deskriptive Psychologie*, Felix Meiner Verlag GmbH, Hamburg; Eng. tr. by Benito Müller 1995, *Descriptive Psychology*, Routledge, London-New York.

Cobb-Stevens, Richard, 2003, "Husserl's Fifth Logical Investigation", in Dahlstrom 2003: 95-107.

Costa, Vincenzo, 1994, "Lo sviluppo della riduzione fenomenologica: dalla *Filosofia dell'aritmetica* a *Ideen*", in *Rivista di Filosofia Neo-Scolastica*, 86, 3: 506-572.

Dahlstrom, Daniel, ed., 2003, *Husserl's Logical Investigations*, Kluwer, Dordrecht.

Derrida, Jacques, 1990, *Le problème de la genèse dans la philosophie de Husserl*, PUF, Paris.

de Warren, Nicolas, 2003, "The Rediscovery of Immanence: Remarks on the Appendix to the *Logical Investigations*", in Dahlstrom 2003: 147-166.

—, 2009, *Husserl and the promise of time: subjectivity in transcendental phenomenology*, CUP, Cambridge (UK).

—, 2018, "The Fountainhead for Brentano's Proteraesthesis and Original Association in Aristotle's Concept of Time", in Moutsopoulos, Evanghelos and Protopapas-Marneli, Maria, eds., *Aristotelis: diachronikòs kai epistīmonikōs epíkairos*, Akadīmia Athīnōn: Kentron Ereunīs tīs Ellīnikīs Filosofias, Athīnai: 247-286.

Fugali, Edoardo, 2005, *Il tempo del sé e il tempo dell'essere. Coscienza e senso interno in Franz Brentano*, La Città del Sole, Napoli.

Husserl, Edmund, 1966, *Zur Phänomenologie des inneren Zeitbewußtseins (1893-1917)*, M. Nijhoff, Den Haag; Eng. tr. by John B. Brough 1991, *On the Phenomenology of the Consciousness of Internal Time*, Kluwer, Dordrecht.

—, 1975, *Logische Untersuchungen. Erster Teil. Prolegomena zur reinen Logik*, M. Nijhoff, Den Haag; Eng. tr. by John N. Findlay 2001, *Logical Investigations*, Vol. 1, Routledge, London-New York.

—, 1976, *Ideen zu einer reinen Phänomenlogie und phänomenlogischen Philosophie. Erstes Buch: Allgemeine Einführung in die reine Phänomenologie*, M. Nijhoff, Den Haag; Eng. tr. by Fred Kersten 1983, *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy. First book. General introduction to a pure phenomenology*, M. Nijhoff, Den Haag.

—, 1979, *Aufsätze und Rezensionen (1890-1910)*, M. Nijhoff, Den Haag; partial Eng. tr. by Dallas Willard 1994, *Early Writings in the Philosophy of Logic and Mathematics*, Kluwer, Dordrecht.

—, 1984, *Logische Untersuchungen. Zweiter Teil. Untersuchungen zur Phänomenologie und Theorie der Erkenntnis. In zwei Bänden*, M. Nijhoff, Den Haag; Eng. tr. by John

N. Findlay 2001, *Logical Investigations*, Voll. 1-2, Routledge, London-New York.

—, 2001, *Allgemeine Erkenntnistheorie. Vorlesung 1902/03*, Kluwer, Dordrecht.

—, 2002a, *Logische Untersuchungen. Ergänzungsband. Erster Teil. Entwürfe zur Umarbeitung der VI. Untersuchung und zur Vorrede für die Neuauflage der Logischen Untersuchungen (Sommer 1913)*, Kluwer, Dordrecht 2002; Eng. tr. by Philip J. Bossert, Curtis H. Peters 1975, *Introduction to the Logical Investigations. A Draft of a Preface to the* Logical Investigations, M. Nijhoff, Den Haag.

—, 2002b, *Zur phänomenologischen Reduktion. Texte aus dem Nachlass (1926-1935)*, Kluwer, Dordrecht.

—, 2003, *Transzendentaler Idealismus. Texte aus dem Nachlass (1908-1921)*, Kluwer, Dordrecht.

—, 2004, *Wahrnehmung und Aufmerksamkeit. Texte aus dem Nachlass (1893-1912)*, Springer, Dordrecht.

Kraus, Oskar, 1919, *Franz Brentano. Zur Kenntnis seines Lebens und seiner Lehre. Mit Beiträgen von Carl Stumpf und Edmund Husserl*, Beck, München.

—, 1930, "Zur Phänomenognosie des Zeitbewußtsein", in *Archiv für die gesamte Psychologie*, 75: 1-22.

Lange, Friedrich Albert, 1866, *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*, J. Baedeker, Iserlohn.

Lavigne, Jean-François, 2005, *Husserl et la naissance de la phénoménologie (1900-1913)*, PUF, Paris.

Lohmar, Dieter, 2012, "Zur Vorgeschichte der transzendentalen Reduktion in den *Logischen Untersuchungen*. Die unbekannte 'Reduktion auf den reellen Bestand'", in *Husserl Studies*, 28: 1-24.

McDonnell, Cyril, 2011, "Husserl's Critique of Brentano's Doctrine of Inner Perception and its Significance for Understanding Husserl's Method in Phenomenology", in *Maynooth Philosophical Papers*, 6: 74–111.

Melandri, Enzo, 1960, "I paradossi dell'infinito nell'orizzonte fenomenologico", in Paci, Enzo, ed., *Omaggio a Husserl*, Il Saggiatore, Milano: 83-120.

Nobili, Filippo, 2022, *La prospettiva del tempo. L'idealismo fenomenologico di Husserl come autoesplicitazione della soggettività trascendentale*, Mimesis, Milano-Udine.

Obstfeld, Kaila, 1983 "The Conflict between Intentionality and Inner Observation", in *Philosophy and Phenomenological Research*, 44, 2: 271-278.

Rinofner-Kreidl, Sonja, 1995/96, "Zeitbewußtsein, innere Wahrnehmung und Reproduktion. Die phänomenologische Zeitlehre in der Auseinandersetzung Husserl – Brentano", in *Brentano-Studien*, 6: 193-227.

Twardowski, Kazimierz, 1894, *Zur Lehre vom Inhalt und Gegenstand der Vorstellung. Eine psychologische Untersuchung*, Alfred Hölder, Wien.

Volpi, Franco, 1987, "Il problema della coscienza del tempo in Brentano", in Mucciarelli, Giuseppe, ed., *Vittorio Benussi nella storia della psicologia italiana*, Pitagora, Bologna: 65-104.

Zahavi, Dan, 1998, "Brentano and Husserl on Self-Awareness", *Études Phénoménologique*, 14, 27/28: 127-168.

—, 1999, *Self-Awareness and Alterity. A Phenomenological Investigation*, Northwestern University Press, Evanston (IL).

—, 2002, "The Three Concepts of Consciousness in *Logische Untersuchungen*", in *Husserl Studies*, 18: 51-64.

Can nonexistent mathematical objects make a difference? Meinongianism, indispensability argument and mathematical entanglement

Simone Cuconato

Abstract: One of the arguments that most influence the debate on the existence of mathematical objects is undoubtedly the indispensability argument. Central to this argument is the Quinean ontic thesis that we are committed to the existence of all the entities we (indispensably) quantify over in our best scientific theories. But what if a different meta-ontological paradigm is adopted? In this paper, I propose a Meinongian interpretation of the indispensability argument. The new reading of the indispensability argument, in accordance with heavy duty platonism, allows me to introduce a new notion of metaphysical dependence that goes by the name of *mathematical entanglement*, and to conclude that nonexistent mathematical objects make a difference to the concrete, physical world.

Keywords: Nonexistent Objects, Mathematical Objects, Meinongianism, Indispensability Argument, Mathematical Entanglement

Introduction

Are there any mathematical objects, such as numbers, sets, or functions? Platonists say "yes", nominalists say "no". According to Alan Baker (2003: 263), the dispute over the existence or nonexistence of mathematical objects does not stem "from any explicit thesis of platonism but from a certain background picture" inherited from traditional platonism. This picture portrays two separate worlds – the concrete world of physical objects and the abstract world of mathematical objects. Again according to Baker, this picture has directed philosophical literature to the claim that the existence of mathematical objects *makes no difference* to the concrete world:

> Since mathematical objects are acausal, the existence or non-existence of mathematical objects makes no difference to the actual arrangement of concrete objects (Cornwell 1992: 80).

That conception, as Bernard Linsky and Edward Zalta (1995) have pointed out, has led many philosophers to conceive of abstract objects *on the model* of concrete objects. Alongside this error, an arguably even more serious one has been made: the indispensability argument (IA) was found to be in harmony

philinq XII, 1-2024, pp. 61-78
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.438
Submitted: 10 October 2022
Accepted: 7 February 2024

with one meta-ontological picture, i.e. Quinean meta-ontology. This error has steered IA towards a "flat"[1] image of ontology and, even more seriously, has made it so difficult to make progress in the debate over IA. In this paper, I will show how it is possible to apply a meta-ontological picture other than the Quinean one to the indispensability argument, i.e. Meinongian meta-ontology. What will result is that: *i*) the real merit of IA is not to establish whether mathematical objects exist, but to point out that mathematical objects make a metaphysical contribution to our best scientific theories; *ii*) our best scientific theories cannot be *mathematical-objects-free; iii*) there is a *mathematical entanglement* within our best scientific theories; *iv*) nonexistent mathematical objects *make a difference* to the concrete world.

Section 1. introduces and discusses Meinongian meta-ontology and, in particular, *strong Meinongianism*. Section 2. proposes a Meinongian reading of IA. Section 3. introduces the notion of *mathematical entanglement* to oppose the idea that the existence of mathematical objects "makes no difference" to the concrete world. Section 4. stresses that the notion of mathematical entanglement is in accordance with a recent form of Platonism that goes by the name of heavy duty platonism (HDP), and also, HDP is in line with Priest's Meinongianism vision of the applicability of mathematics. Section 5. concludes by focusing on the fact that mathematical objects *make a difference* to the concrete world even if they are nonexistent objects.

1. *Meinongian meta-ontology and strong Meinongianism*

The debate between Quineans and Meinongians has a long history, and most of its key episodes antedate the recent, burgeoning development of *meta-ontology*.[2] The mainstream Quinean meta-ontology claim to what has been called "the question of ontology" is:

(Q) Everything exists, because it makes no sense to speak of "nonexistent objects".

That is the apparently *easy* and *tautological* reply to that question: "What exists?". In his introduction to ontology Achille Varzi writes:

[1] The attribute "flat" is used by Jonathan Schaffer to indicate the structure of Quinean metaphysics: "For the Quinean, the target is flat. The task is to solve for E = the set (or class, or plurality) of entities. There is no structure to E. For any alleged entity, the flat conception offers two classificatory options: either the entity is in E, or not" (Schaffer 2009: 354).

[2] The term "meta-ontology" was introduced by Peter van Inwagen (1998). For van Inwagen the key question for meta-ontology is "What do we mean when we ask "What is there", and "What is the correct methodology of ontology?". A comprehensive introduction to the topic is (Berto, Plebani 2015).

As Quine has written, everything exists because it makes no sense to speak of "nonexistent entities", and those who think otherwise would manifest, not an ontological disagreement, but a misunderstanding of the very concept of existence. […] Precisely because it would be inconsistent to claim that something does not exist, however, to claim that everything exists is tautological, that is, devoid of content, therefore of interest (Varzi 2005: 3, my translation).

On the contrary for a Meinongian, pace Quine, there are objects that *do not exist*:

(M) There are things that do not exist.

Meinongianism was originally defended by Alexius Meinong (1904), and the view has recently been developed and defended by Terence Parsons (1980), Richard Routley (1980), Edward Zalta (1988), Graham Priest (2005) and Francesco Berto (2013; 2015). The order of the names above is not casual but corresponds to a precise way of understanding the Meinongian *Principle of Comprehension*. Indeed, any Meinongian theory needs some principle stating which objects are admitted by the theory, and which properties they can have. In its naïve version Meinongianism defends to what Parsons (1980) called an "Unrestricted Comprehension Principle" (UCP) for objects:

(UCP) For any condition $A[x]$, with free x, some object satisfies $A[x]$.

UCP is the foundation of the naïve version of Meinongianism, which holds that for every given set of properties there is some object that exactly bears those properties. Last century, this view was widely criticised by Bertrand Russell (1903; 1905a; 1905b) and by Willard Van Orman Quine (1948) on three several counts:[3]

1) the property of existence[4] can be added to any combination of properties to yield existent objects that are incomplete (e.g. an existent object that is green and nothing else);
2) the existence of inconsistent object (e.g. round square), or that straightforwardly disagrees with empirical observation (e.g. golden mountain on Earth);
3) it is not clear how an adequate criterion of identity can be given for the domain of nonexistent objects.

[3] On why there is no contradiction in the Meinongian thesis and why the Quinean thesis is not clear see (Sendłak 2023).

[4] Of course, it only makes sense to talk of "properties of existence" from a Meinongian point of view or for those who defend a *classical* position of existence and essence. From Frege to Quine and for many philosophers today, existence is not a property. For an analysis of the limitations of Frege-Quine's quantificational theory of existence see (Galvan 2015), while for a new system of logic for existence and essence see (Galvan, Giordani 2020).

Since many of the objections to Meinongianism have focused on the Unrestricted Comprehension Principle, in the last last few decades different approaches to the problem have been developed. One approach, Nuclear Meinongians, limited UCP to a restricted vocabulary and distinguished between two kinds of predicates, called *nuclear* and *extranuclear* (Parsons 1980; Routley 1980). The second approach, by dual copula Meinongians, distinguishes between two ways in which things can be ascribed properties: ordinary predication expressing property-instantiation and *encoding* (Zalta 1983; 1988). Finally, the third approach is far more recent, to be called Modal Meinongianism, and claims that an object has all the properties it is characterized as having not necessarily in the actual world but in some possible or impossible worlds (Priest 2003; 2005; Berto 2013; 2015; Berto, Priest 2014). Regardless of which approach of neo-Meinongianism one wishes to defend, one can correctly say that Meinongianism is understood as "the view that some objects do not exist, but we can generally talk about them, quantify on them, and state true things about them" (Berto 2013: 115). I agree with Filippo Casati and Naoya Fujikawa that, given this definition, contemporary Meinongians support the following three ideas:

[1] Some objects do not exist.
[2] We can refer to and quantify over nonexistent objects to make true statements about them.
[3] Nonexistent objects are (or, at least, are parts of) the truth-makers of sentences which contain reference to or/and quantification over nonexistent objects.

Casati and Fujikawa call a position which endorses [1], [2] and [3] *strong Meinongianism*:

> And contemporary Meinongians like Parsons, Routley, Zalta and Priest seem to take the property-possession of nonexistent objects as contributing to truth-making of sentences about nonexistent objects. For example, even if a is a nonexistent object, if a has the property P, then, the sentence 'a is P' is true simply because of a's possession of P. In this way, nonexistent objects can be involved in truth-making of sentences about them. Moreover, Meinongianism accepts that a nonexistent object can have properties which are not parts of its characterization. […] Because of this, according to these contemporary Meinongians, nonexistent objects can be involved in the truth-makers of sentences such as 'Sherlock Holmes is possible' and so on. Let us call a position which endorses [1], [2] and [3] strong Meinongianism. Many contemporary Meinongians seem to endorse strong Meinongianism (Casati, Fujikawa: 424-425).

The strength of this position, which will play a key role in this paper, is that it overcomes some of the problems of *reductionist approach* regarding at the instantiation of the property of being possible by nonexistent objects, and "the tension between admitting quantification over nonexistent objects, on the one hand, and the reductionist approach which tries to eliminate nonexistent objects from truth-making, on the other hand" (*ivi*: 433). The solution, according to Casati and Fujikawa, is to adopt strong Meinongianism by admitting truth-making by nonexistent objects.

2. *The indispensability argument*

Generally, the indispensability argument was understood as an argument for the existence of abstract mathematical objects (Colyvan 2001; Panza, Sereni 2013; Cuconato 2022). For those who support IA we should believe in mathematical objects for the same reason as we believe in microphysical particles: because our best scientific explanations imply their existence. In the words of Hartry Field:

> [I]f our belief in electrons and neutrinos is justified by something like inference to the best explanation, isn't our belief in numbers and functions and other mathematical entities equally justified by the same methodology? After all, the theories that we use in explaining various facts about the physical world not only involve a commitment to electrons and neutrinos, they involve a commitment to numbers and functions and the like (Field 1989: 16).

This suggests that the most direct formulation of the argument should be something like this (Liggins 2016: 532):

i) Mathematics is indispensable to science: that is, our best explanations imply the existence of numbers and other mathematical objects;

ii) If mathematics is indispensable to science, then there are mathematical objects;

[I] ———————————

iii) There are mathematical objects.

Another of the most debated formulations of IA has been advanced by Mark Colyvan (2001: 11):

i) We ought to have ontological commitment to all and only those entities that are indispensable to our best scientific theories;

ii) Mathematical entities are indispensable to our best scientific theories;

[II] ———————————

iii) We ought to have ontological commitment to mathematical entities.

According to Otávio Bueno "the indispensability argument crucially relies on the conditions under which quantification over certain objects is indispensable" (Bueno 2018: 203). There are three conditions identified by Bueno (*ivi*: 203-204) and they correspond to three distinct roles played by mathematical objects in scientific theories:

a) Expressive role: Quantification over certain objects is, in many cases, made in order to express certain facts (situations, possibilities, etc.). If such expressions cannot be formulated without reference to the objects in question, the quantification is expressively indispensable.
b) Predictive role: Objects are often quantified over in predictive contexts in which they are invoked to make predictions about certain phenomena. If such predictions cannot be obtained without reference to the relevant objects, the quantification is predictively indispensable.
c) Explanatory role: Objects are frequently quantified over to explain certain phenomena. If the explanations in question cannot be implemented without reference to such objects, the quantification is explanatorily indispensable.

In turn, the conditions under which quantification over certain objects is indispensable are based on the standard view of existence, i.e. Quine's criterion for ontological commitment. I will call these forms of IA *standard indispensability argument* (S-IA). Standard because the standard view of existence is applied, i.e. Quine's criterion for ontological commitment. Generalizing, Quine's criterion of ontological commitment can be formulated as follows:

(QC) An object *o* is an ontological commitment of a regimented (set of/) statement(s) iff *o* is required to make the statement(s) true (where *o* is assumed to be in the range of the bound variable(s)).

Quine puts it this way: "To be assumed as an entity is…to be reckoned as the value of a variable", and "we look to bound variables in connection with ontology not in order to know what there is, but in order to know what a given remark or doctrine…says there is" (Quine 1948: 13). A contemporary Quinean is Peter Van Inwagen (1998). Van Inwagen sums up Quine's position in a series of five theses:

1. Being is not an activity.
2. Being is the same as existence.
3. Existence is univocal.
4. Existence is expressed by the existential quantifier.

5. (QC) is a procedural norm for ontological disputes.

Let's focus on thesis 5.[5] According to Ted Parent (2019) "(QC) lays down one of the 'rules for engagement' for ontological disputes".[6] It is precisely this rule for engagement that has tied the debate on the subject of indispensability almost exclusively to Quinean meta-ontology. In this sense, it's fair to say that QC is the *indispensable meta-ontological premise* to the standard version of IA. But what if we read the indispensability argument through the glasses of a Strong Meinongian? The structure of IA would not change. What would totally change is the way we understand that *there are mathematical objects*. For a Meinongian to say that there are objects of type does not mean that objects of type exist. Not only that, if we accept the position of the majority of neo-Meinongians as Parsons, Routley, Zalta, Priest or Berto, mathematical objects are abstract objects and "abstracts objects are just kind of non-existent object" (Priest 2005: 135):[7]

[4] Mathematical objects are nonexistent objects.

If we replace QC with with [1], [2], [3], [4] the most direct formulation of IA should be something like this:

i) Mathematics is indispensable to science: that is, our best explanations imply that *there are* numbers and other mathematical objects;
ii) If mathematics is indispensable to science, then *there are* mathematical objects;

[III] ———————————

iii) *There are* nonexistent mathematical objects.

iv) Nonexistent mathematical objects are the truth-makers of sentences which contain reference to or/and quantification over mathematical objects.

I will call this form of IA *non-standard indispensability argument* (NS-IA). Non-standard because expressions such as "there are" must be understood in the sense of "Meinongian quantification". To say that "there are nonexistent mathematical objects" is not to say that such objects exist. In this way, the focus is shifted from "rules for engagement-existence" to "nonexistent objects-truthmaker". The argument of indispensability thus developed without

[5] More precisely thesis 5 is not given a one-sentence formulation in van Inwagen; but he says that it is really a "family of theses" (2014: 85).

[6] Not by chance, van Inwagen illustrates this using the Platonism vs. nominalism debate about numbers.

[7] It is interesting to note that for Meinong (1904) abstract objects do not exist, but they do *subsist*.

denying the conditions under which quantification over certain objects is indispensable, allows us to conclude that "there are" mathematical objects, but we are not ontologically committed to them – at least in the Quinean sense of ontological commitment. What I want to argue is that the cause of the stalemate in the IA debate was precisely Quine's quantificational theory of existence. In this way, the debate was regimented on the topic "*rules for engagement-existence-indispensability*".

Also [III], like [I] and [II], is based on the three conditions under which quantification over certain objects is indispensable, but in NS-IA changes the way quantification over mathematical objects is understood to express, predict, and explain a variety of features of the world. If Quine "used the indispensability argument as a reason to support ontological commitment to classes" (Bueno 2018: 204), in contrast, a meta-ontological perspective such as Meinongianism, allows us to consider IA valid even if mathematical objects do not exist and to use IA as a reason to support the metaphysical role that mathematical objects play in our best scientific theories. NS-IA admits truth-making by nonexistent objects, and therefore, the strength of analysis of NS-IA, is to shift the focus to the fact that mathematical objects *make a difference* to the concrete world even if they are nonexistent objects.

3. *Orthogonality and mathematical entanglement*

In recent years, the scientific literature on IA has developed a critique inspired by the idea that the existence of mathematical objects "makes no difference" (MND) to the concrete, physical world. Baker (2003) provides effective critiques of some ways in which "makes no difference" claims have been defended. I totally agree with Baker (2003) that there is a tight connection between IA and "makes no difference" claims, however, I prefer Matteo Plebani's formulation of MND. Here is how Plebani (2018: 256) proposes to understand MND claims:

– (Orthogonality) The two subject matters how the concrete world is and whether there are abstract objects are orthogonal.

In detail, Plebani referring to Lewis' original definition of subject matter, and the works of Stephen Yablo (2012; 2014), understands orthogonality "as the view that the way the concrete world is does not demand nor preclude the existence of abstract mathematical objects (Yablo 2012): given a world in which mathematical objects are present (absent) there is a world *w'* in which mathematical objects are absent (present) and that is concretely indiscernible from *w"* (Plebani 2018: 257).

Nominalists, relying on the notion of causal dependence, have been attracted by orthogonality:

– (BAD-ProO) Mathematical entities have no causal powers; therefore the existence of mathematical entities makes no difference for how the concrete world is.

Nevertheless, as Baker (2003) points out, this argument for orthogonality is unsatisfactory because (BAD-ProO) is fundamentally based on a certain background picture inherited from traditional platonism: mathematical objects have no causal powers and therefore "mathematical objects make no difference for the concrete world from the fact that mathematical objects make no difference for the concrete world at the causal level" (*ivi*: 257). Those who defend (BAD-ProO) do not take into account that there are other forms of dependence besides causal dependency, e.g. metaphysical dependence.

Alongside (BAD-ProO) Plebani also presents a bad argument developed by Yablo:

– (BAD-AntiO) "we cannot imagine-without-numbers a complex world" therefore "we cannot imagine a complex world lacking in numbers" (Yablo 2012: 1014).

Those who defend (BAD-AntiO) write Plebani "confuses the issue of what it takes to describe certain circumstances with the issue of what it takes for those circumstances to obtain" (Plebani 2018: 257). In general, I agree that (BAD-AntiO) is not such a strong argument (although I find the issue of "what it takes" quite obscure). However, what I am interested in proposing now is not a criticism or a defence of the two positions mentioned above, but rather to present a new argument against orthogonality.

One of the key points of my paper is the distinction between an MND developed from an *ontological* point of view and an MND developed from a *metaphysical* point of view. The distinction between the ontological and metaphysical planes is the result of the Meinongian thesis that some things, that is, some property-bearers or *Sosein*-bearers, are nonexistent or lacking *Sein*. This thesis is the famous Meinongian *Principle of Independence* states that an object's *Sein*, its existential status, is independent from its *Sosein*, its having properties.[8] Given

[8] It is important to specify that the Principle of Independence: *i*) is connected by Meinong to the Principle of the *Aussersein*, or "principle of the indifference of pure Objects to being" (Meinong 1904: 86): being and non-being are not part of the Sosein of the thing; and *ii*) is set against a very strong metaphysical thesis: called *serious actualism*, consisting in the thesis that the having of whatever property entails existence. Serious actualism has deep roots in the history of philosophy (Berto 2013: 86) but it is especially in the actualism of Russell the idea that the having of any property entails existence. Quinean meta-ontology was influenced by Russellian actualism and, consequently, serious actualism directed IA and its subsequent debate. Finally, it is historically and theoretically important

the Principle of Independence, to say that an object *X* has the property of makes a difference to the concrete world, does not also mean that objects *X* exist.

In the previous section, I highlighted the weight that the choice of a particular meta-ontological picture occupies in IA. In particular, I interpret Quinean "regimentation" as the cause of the stalemate in the IA debate: we translate a scientific theory into classical first-order logic, then read off its ontological commitments as per (QC), we will have a criterion for establishing what objects such a theory says exist. In a Meinongian perspective IA does not establish that mathematical objects exist, but rather that mathematical objects contribute to the truth-making of sentences concerning concrete objects. Therefore, given [III], it is possible to reformulate MND by placing the emphasis no longer on the ontological aspect (there are abstract objects) but on the metaphysical one:

– (New Orthogonality) The two subject matters how the concrete world is and whether there are abstract objects that make a metaphysical contribution to our best scientific theories are orthogonal.

The real merit of the indispensability argument and "makes no difference" claims is not to establish whether mathematical objects exist or not exist, but whether mathematical objects make a metaphysical contribution to our best scientific theories. In this sense, [III] brings out the indispensability of mathematical objects in the sense that our best scientific theories cannot be *mathematical-objects-free* because: *i*) mathematics plays a rolee expressive, inferential, and explanatory in the scientific practice and; *ii*) mathematical objects contribute to the truth-making of sentences concerning concrete objects. In other words, to say that our best scientific theories cannot be mathematical-objects-free is to say that there is an inextricable *mathematical entanglement* within our best scientific theories. The term "entanglement" is inspired by *quantum entanglement* and was coined by Erwin Schrödinger in 1935. John Clauser, Anton Zeilinger and Alain Aspect have won the 2022 Nobel Prize in Physics for their experiments with entangled photons, in which particles

to specify that the philosophical roots of the principle of independence are well established in Scholastic metaphysics in the doctrine of the *real distinction* between the *essence* of the thing (that is to say, what the thing is) and the *existence* of the thing (that it is) in authors such as Avicenna, Duns Scotus and especially Thomas Aquinas. As Edward Feser points out: "A third argument (with Aquinas presents in Chapter IV of *On Being and Essence*) holds that we can know the essence of a thing without knowing one way or the other whether it exists. Suppose a person had never before heard of lions, velociraptors, or unicorns, and you give him a thorough description of the nature of each. You then tell him that of these three creatures, one exists, one used to exist but is now extinct, and the third never existed; and you ask him to tell you which is which given what he now knows about their essences. He would, of course, be unable to do so. But then existence of the creatures that exist must be really distinct from their essences, otherwise one could know of their existence merely from knowing their essences" (Feser 2014: 243).

of light become *inextricably linked*.[9] In the same way as quantum particles, physical objects are inextricably linked to mathematical objects. Because of the effects of mathematical entanglement, the properties of an abstract mathematical object influence the properties of a concrete physical objects. This mathematical entanglement makes it clear that the relations between abstracta and concrete are metaphysically fundamental and they do not hold *in virtue of* any nominalistic properties possessed by the concrete objects alone.[10] Based on this, it is possible to formulate a good argument against (new) orthogonality:

– (GOOD-AntiO): since mathematics plays a rolee expressive, inferential, and explanatory in scientific theorizing and mathematical objects contribute to the truth-making of sentences concerning concrete objects, our best scientific theories cannot be mathematical-objects-free; therefore, our best scientific theories are committed to *mathematical entanglement*; the relation between abstracta and concreta are metaphysically fundamental; and mathematical objects "make a difference" to the concrete, physical world.

In this way, on the one hand, Meinongian meta-ontology makes it possible to avoid commitment to mathematical objects despite granting that quantification over them is indeed indispensable to our best theories of the world, on the other hand, (GOOD-AntiO) shows that the concrete world demands abstract mathematical objects.[11] This seemingly contradictory *phenomenon* is possible because mathematical objects should not be classified as pure *fictional objects*:[12] "things mentioned and described in tales, novels, fantasy stories, operas, and so on" (Berto, Plebani 2015: 104). Indeed, while it is true that existence should not be required for successful reference or quantification to objects that do not exist,[13] such as Sherlock Holmes or the number 2, it is also true that the relationship that our scientific theories have with nonexistent objects is profoundly

[9] This phenomenon leads to paradox and has puzzled physicists for a long time because the behavior of entangled particles is apparently inexplicable and incomprehensible. The trio's experiments proved that connections between quantum particle comes from a genuine association in which manipulating one quantum object affects another far away. Albert Einstein famously called the phenomenon "spooky action at a distance" – it is now known as quantum entanglement.

[10] In this sense, the notion of "mathematical entanglement" is related to the notion of "grounding". The *grounding theory* refers to a relation of non-causal dependence, or metaphysical dependence, generally introduced by the use of terms like "in virtue of" or "because". See Liggins (2016) for a nominalist strategy developed through the connection between grounding and IA.

[11] An argument from the presumption that mathematics is indispensable for science to the claim that mathematical objects make a difference to the concrete world is developed by Colin Cheyne and Charles Pigden (1996).

[12] In general, for a study of the advantages of Meionongianism over fictionalism see (Gan 2021).

[13] Bueno (2018: 215) argues in detail how it is possible to successfully refer to objects that do not exist or successfully quantifies over objects that do not exist.

different depending on the *type* of nonexistent object. Consider, for instance, complex numbers. Generally, the number system we all know, like 2, -5, 0.8, √2, 16/4, and π, are known as the real numbers ℝ, while, a complex number is a number written in the form:

$$z = a + bi$$

where *a* and *b* are real numbers.[14]

Complex numbers have many applications in physics, chemistry, biology, electrical engineering, statistics, and economics. In particular, they have become fundamental in quantum mechanics:

> Quantum mechanics seems to use complex numbers in a more fundamental way. It suffices to look at some of the most basic equations, both in the matrix $([\hat{p}, \hat{x}]) = ih$ and wave $((ih(\partial\psi/\partial t = \hat{H}\psi))$ formulations, to wonder about the presence of the imaginary unit. This complex nature of quantum mechanical quantities puzzled some of the very founders of the theory (Karam 2020: 39).

It becomes clear that complex numbers play a fundamental role in scientific practice and contribute to the truth-making of sentences concerning concrete objects. In contrast, for other nonexistent objects, such as Sherlock Holmes, Harry Potter or Alice in Wonderland, their metaphysical role and contribution in scientific theorizing cannot be traced.[15] Therefore, complex numbers are inextricably linked to physical objects and make a metaphysical difference to the way the concrete world is. One moral to be drawn from (GOOD-AntiO) is that the concrete world demands abstract mathematical objects, and the relations between abstracta and concrete are not *grounded* in the nominalistic properties of concrete objects.[16]

4. *Heavy duty platonism*

The last aspect I want to address is how the concept of mathematical entanglement is inherent in a recent form of Platonism that goes by the name of *heavy duty platonism* (HDP). Also, HDP is in line with Priest's Meinongianism vision of the *applicability of mathematics*. Let's proceed in order. HDP is a view about the metaphysics of physical magnitudes, like temperature and mass. Robert Knowles characterizes HDP as "the view that physical magnitudes, such as mass and tem-

[14] *a* is known as the "real part" of *z*, and *b* as the "imaginary part". Furthermore, we define the imaginary unit number *i* as the square root of -1, that is: $i = \sqrt{(-1)}$.

[15] In this way, the metaphysical structure of nonexistent objects is not *flat*.

[16] This aspect differentiates my position from mathematical fictionalism.

perature, are cases of physical objects being related to numbers" (Knowles 2015: 1255). The key point is how HDP understands this "being related":

> what they tell us about these relations: weaker forms tell us they are derivative of more fundamental properties or relations that hold of physical objects alone, while HDP says these relations are fundamental (*ivi*: 1256).

One of the most interesting arguments formulated by Knowles is based on the Lewisian distinction between intrinsic and extrinsic properties:

> A thing has its intrinsic properties in virtue of the way that thing itself, and nothing else, is. Not so for extrinsic properties, though a thing may well have these in virtue of the way some larger whole is. The intrinsic properties of something depend only on that thing; whereas the extrinsic properties of something may depend, wholly or partly, on something else (Lewis 1983: 111-112).

Knowles attacks the intuition that some physical magnitude properties (such as mass) are clear-cut cases of intrinsic properties. Those who defend this insight believe that an object's mass involves only that object; and since HDP implies that an object has its mass by being related to something external such as a number, HDP implies mass is extrinsic and, therefore, HDP is false. One of the arguments in defence of HDP is to show how the intuitions we have about the properties of physical phenomena are unreliable. In the Standard Model – our best physical theory of fundamental particles – the mechanism by which a particle gains its mass is based on the assumption of the existence of a scalar field entirely distinct from any particle, the Higgs Field:

> The mass of particles depends not only on the properties of the particles, but also on the properties of the Higgs Field. The Higgs Field is entirely distinct from any particle interacting with it. For any particle *p*, there is a world in which *p* does not exist while the Higgs Field does. Science tells us that the mass of a particle depends on the properties of something distinct from the particle, so, on the present view, mass is extrinsic. According to HDP, then, mass is a three-place relation holding between an object, a number, and the Higgs Field (Knowles 2015: 1260).

This fundamental form of the relationship between abstracta and concrete is precisely expressed by *mathematical entanglement*. To say that there is a mathematical entanglement within our best scientific theories, means to say from the point of view of HDP "that there is a robust metaphysical connection between physical objects and numbers that renders the latter explanatorily relevant to physical phenomena" (*ibidem*). It is the mathematical entanglement that offers a possible explanation for why such abstract objects as complex numbers make a difference to the concrete world:

Complex numbers seem to be fundamental for the description of the world proposed by quantum mechanics. In principle, this can be a source of puzzlement: Why do we need such abstract entities to describe real things? (Karam 2020: 43).

We need the complex numbers to describe "real things" because there is a solid metaphysical connection between physical objects and mathematical objects, that is, between existent objects and nonexistent objects, that is, between causal entities and non-causal entities.

But how can mathematical entanglement and HDP help in the question of the applicability of mathematics? Priest (2003) addresses the problem by asking the question from a Meinongian point of view: "How can non-existent objects tell us anything about existent ones?" (*ivi*: 11). Priest after referring to Routley's proposed solution goes straight to the heart of the problem:

How, then, is one to explain the fact that properties of non-existent objects can tell us something about existent objects? Actually, exactly the same question can be posed for platonism, and the answer in both cases is the same. The physical quantities in question have certain properties, and the mathematical quantities have other properties. But we can move between the one and the other because these properties have the same structure, and, specifically, because the correlation established by the bridge laws is an isomorphism (*ivi*: 12).

According to Priest we can move between physical and mathematical properties because these properties have the same *structure*.[17] And this structure is exactly the robust metaphysical connection between concrete physical objects and mathematical abstract objects that HDP subscribes to. In this way, the fact that the properties of nonexistent objects can tell us something about existent objects is explained by the fact that existent objects being metaphysically related to nonexistent objects. And that is why our best scientific theories explaining how the concrete world is made are instances of mathematical entanglement. For this, Priest's meinongianism differs from weak versions of platonism because according to HDP platonic relations between nonexistent objects and existent objects are metaphysically fundamental. In the words of Knowles:

It is quite plausible to hold that some physical objects have some of their causal powers by instantiating relations to non-causal entities (Knowles 2015: 1267).

Within this framework, at least two fundamental questions remain unanswered and will have to be investigated in the future. The first, concerns the form of mathematical structuralism that might be in accordance with NS-IA

[17] Mathematical structuralism is the view that pure mathematics is about abstract structure or structures.

and HDP. In general, I think one must investigate those forms of structuralism that "acknowledges that abstract structures exist, that the pure objects of mathematics are in some sense elements of, or places in, those structures, and that there is nothing more to the pure objects of mathematics than can be described by the basic relations of their corresponding structure" (Nodelman, Zalta: 39).[18] The second, concerns the possibility of defining a rigorous metaphysical justification for mathematical entanglement within the model of meta-metaphysics developed by Alessandro Giordani and Ciro De Florio (2020).[19]

5. *Conclusion*

The present paper has shown the following: first, the real merit of IA is not to establish whether mathematical objects exist, but to point out that mathematical objects make a difference to the concrete world. Secondly, our best scientific theories cannot be mathematical-objects-free is because there is a mathematical entanglement within our best scientific theories. Thirdly, the fact that the properties of nonexistent objects can tell us something about existent objects is explained by the fact that existent objects being metaphysically related to nonexistent objects. Therefore, the strength of NS-IA lies in shifting the focus to the statement that mathematical objects make a difference to the concrete world even if they are nonexistent objects.

Simone Cuconato
University of Calabria
simone.cuconato@unical.it

Acknowledgements: I would like to express my sincere gratitude to Sergio Galvan and Annabella d'Atri for their invaluable guidance and mentorship in logic and metaphysics. I would also like to thank the anonymous referees of this journal for their highly valuable comments on a previous version of this work.

[18] On this form of structuralism see (Dedekind 1888; Parsons 1990; Shapiro 1997) and in particular mathematical structuralism of Uri Nodelman and Zalta. Their structuralism is based on the following insights: "First, […], abstract objects are connected to the properties that define them in a different way than ordinary objects are connected to the properties they bear. Second, […], theorems and truths about abstract relations are more important in defining mathematical structures than mathematical entities" (Nodelman, Zalta 2014: 40).

[19] Giordani and De Florio proposed a new meta-metaphysics model (*Metaphysics Constrained by Physics or Logic*) within which they acknowledge: *i*) an autonomous field of enquiry for metaphysics; and *ii*) the possibility for metaphysics to be based on justification procedures that are "in part different from the empirical procedures of science or the a priori procedures of logic" (De Florio, Giordani 2020: 948, my translation).

References

Baker, Alan, 2003, "Does the existence of mathematical objects make a difference?", in *Australasian Journal of Philosophy*, 81, 2: 246-264.

Berto, Francesco, 2013, *Existence as A Real Property: The Ontology of Meinongianism*, Springer, London.

—, 2015, "There Is an 'Is' in 'There Is': Meinongian Quantification and Existence", in Torza, Alessandro (ed). *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, Synthèse Library, vol 373, Springer, Cham: 221-240.

—, Matteo Plebani, 2015, *Ontology and Metaontology. A Contemporary Guide*, Bloomsbury.

—, Graham Priest, 2014, "Modal Meinongianism and Characterisation: Reply to Kroon", in *Grazer Philosophische Studien*, 90, 1: 183-200.

Bueno, Otavio, 2018, "Putnam's indispensability argument revisited, reassessed, revived", in *Theoria. An International Journal for Theory, History and Foundations of Science*, 33, 2: 201-218.

Casati, Filippo - Naoya Fujikawa, 2016, "Nonexistent Objects as Truth-Makers: Against Crane's Reductionism", in *Philosophia*, 44: 423-434.

Cheyne, Colin, Charles R. Pigden, 1996 "Pythagorean Powers, or A Challenge to Platonism", in *Australasian Journal of Philosophy*, 74, 4: 639-45.

Colyvan, Mark, 2001, *The Indispensability of Mathematics*, Oxford University Press, Oxford-New York.

Cornwell, Stuart, 1992, "Counterfactuals and the Applications of Mathematics", in *Philosophical Studies*, 66: 73-87.

Cuconato, Simone, 2022, *Impegno ontologico e l'argomento di indispensabilità in filosofia della matematica*, Il Sileno Edizioni, Lago (CS).

—, 2024, "Fol-Based Applied Ontology for Metadata Extraction in Mathematical Knowledge Management", in *Romanian Journal of Mathematics and Computer Science*, 14, 1: 1-11.

Dedekind, Richard, 1888, *Was sind und was sollen die Zahlen*, Vieweg und Sohn, Braunschweig.

De Florio, Ciro, Alessandro Giordani, 2020, "Alcune ipotesi sulla possibilità della metafisica", *Rivista di Filosofia Neo-Scolastica*, 4: 943-967.

Feser, Edward, 2014, *Scholastic Metaphysics. A Contemporary Introduction*, editiones scholasticae, Heusenstamm.

Field, Hartry, 1989, *Realism, mathematics and modality*, Blackwell, Oxford.

Galvan, Sergio, 2015, "Quantificatori ed esistenza", in *Rivisita di Filosofia Neo-Scolastica*, 1-2: 143-153.

—, Alessandro Giordani, 2020, "A Classical Logic of Existence and Essence", in *Logic and Logical Philosophy*, 29, 4: 541-570.

Gan, Nathaniel, 2021, "Fictionalism and Meinongianism", in *Theoria. An International Journal for Theory, History and Foundations of Science*, 36, 1: 49-62.

Karam, Ricardo, 2020: "Why are complex numbers needed in quantum mechanics? Some answers for the introductory level", in *American Journal of Physics*, 88: 39-45.

Knowles, Robert, 2015, "Heavy Duty Platonism", in *Erkenntnis*, 80, 6: 1255-1270.

Lewis, David, 1983, "Extrinsic Properties", in *Philosophical Studies*, 44, 197-200.

—, 1988, "Statements partly about observation", in *Philosophical Papers*, 17, 1: 1-31.

Liggins, David, 2016, "Grounding and the indispensability argument", in *Synthese*, 193, 2: 531-548.

Linsky, Bernard, Edward N. Zalta, 1995, "Naturalized Platonism versus Platonized naturalism", in *The Journal of Philosophy*, 92: 525-555.

Meinong, Alexius, 1904, "Über Gegenstandstheorie", Meinong, Alexius (eds). *Untersuchungen zur Gegenstandstheorie und Psychologie*, J.A. Barth, Leipzig: 1-51; Eng. tr. 1960, "The Theory of Objects", in Chisholm, ed., *Realism and the Background of Phenomenology*, The Free press-Collier-Macmillan, New York & London: 76–117.

Nodelman, Uri, Edward N. Zalta, 2014, "Foundations for Mathematical Structuralism", in *Mind*, 123, 489: 39-78.

Panza, Marco, Andrea Sereni, 2013, *Plato's Problem. An Introduction on Mathematical Platonism*, Palgrave Macmillan, London.

Parent, Ted, 2020, "Ontological Commitment and Quantifiers", in R. Bliss, J.T.M. Miller (eds), *The Routledge Handbook of Metametaphysics*, Routledge, Abingdon.

Parsons, Charles, 1990, "The Structuralist View of Mathematical Objects", in *Synthese*, 84: 303-346.

Parsons, Terence, 1980, "*Nonexistent objects*", Yale University Press, New Haven and London.

Plebani, Matteo, 2018, "The indispensability argument and the nature of mathematical objects", in *Theoria: An International Journal for Theory, History and Foundations of Science*, 33, 2: 249-263.

Priest, Graham, 2003, "Meinongianism and the philosophy of mathematics", in *Philosophia Mathematica*, 11, 1: 3-15.

—, 2005, *Towards Non-Being: The Logic and Metaphysics of Intentionality*, Oxford University Press, Oxford.

Quine, Willard Van Orman, 1948, "On What There Is", in *Review of Metaphysics*, 2: 21-38.

Routley, Richard, 1980, *Exploring Meinong's Jungle and Beyond*, Australian National University, RSSS, Canberra.

Russell, Bertrand, 1903, *Principles of Mathematics*, Cambridge University Press, Cambridge.

—, 1905a, "On denoting", in *Mind,* 14: 479-493.

—, 1905b, "Review of A. Meinong, Untersuchungen zur Gegenstandtheorie und Psychologie", in *Mind*, 14: 530-538.

Schaffer, Jonathan, 2009, "On what grounds what", in David Manley, David J. Chalmers, Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford University Press, Oxford: 347-383.

Sendłak, Maciej, 2023, "From the Meinongian Point of View", in *Logic and Logical Philosophy*, 32, 2: 243-251.

Shapiro, Stewart, 1997, *Philosophy of Mathematics: Structure and Ontology*, Oxford University Press, New York.

Van Inwagen, Peter, 1998, "Meta-ontology", in *Erkenntnis*, 48: 233-50.

Varzi, Achille C., 2005, *Ontologia*. Laterza, Rome.

Yablo, Stephen, 2012, "Explanation, extrapolation, and existence", in *Mind*, 121, 484: 1007-1029.

—, 2014, *Aboutness*, Princeton University Press, Princeton.

Zalta, Edward, 1983, *Abstract objects: An introduction to axiomatic metaphysics*, D. Reidel, Dodrecht.

—, 1988, *Intensional Logic and the Metaphysics of Intentionality*. MIT Press, Cambridge, MA.

Focus

Introduction: The role of experts in democratic societies: In honor of Paul K. Feyerabend

Pierluigi Barrotta, Gustavo Cevolani, Roberto Gronda, Luca Tambolo

Paul K. Feyerabend (1924-1994) was a very prolific writer. In the course of a career that spanned almost five decades, he tackled a wide array of issues, ranging from technical problems in the philosophy of science – such as explanation, the role of experience in scientific theorizing, the importance of alternatives in theory testing, and scientific method – to questions of more general interest such as the proper role of scientific experts in decision-making within democratic societies. As Feyerabend's centennial was approaching, in June 2023 the IMT School for Advanced Studies Lucca and the University of Pisa jointly convened a workshop to honor his work, held at the IMT School's campus with the financial support of the Horizon 2020 project Inclusive Science and European Democracies (ISEED) and by the Italian Ministry of Education, University and Research (MUR).[1] The general theme of the workshop, from which the present Focus of *Philosophical Inquiries* stems, was "The Role of Experts in Democratic Societies. In Honor of Paul Feyerabend". However, as readers will readily appreciate, the proceedings covered a wider terrain. This came as no surprise. In fact, Feyerabend's views on expert policy advice, which he expounded mainly in writings published from the late 1960s onwards and to which he devoted, for instance, significant chunks of his book *Science in a Free Society* (1978), were deeply informed by his views on such issues as the importance of theoretical pluralism within scientific inquiry, on which he had been steadily working since the 1950s. The articles included in the present Focus, which we will briefly introduce in what follows, clearly reflect this feature of Feyerabend's philosophical work.

In "Feyerabend, Experts, and Dilettantes", John Preston scrutinizes Feyerabend's paper "Experts in a Free Society" (1970/1999). Comparatively less studied than Feyerabend's most well-known contributions on the issue of sci-

[1] Through the PRIN 2017 Project "From Models to Decisions" (grant nr. 201743F9YE) and the PRO3 Project "Understanding Public Data: Experts, Decisions, Epistemic Values" (grant nr. D67G22000130001).

entific expertise, "Experts in a Free Society" certainly makes for an engaging read. In particular, Feyerabend's attempt to ridicule the writing style of contemporary scientific experts, as exemplified by *Human Sexual Response* (Master *et al.* 1966), looks pretty successful. At first sight, at least. In fact, as Preston shows, Feyerabend's critique of experts relies on a number of very questionable argumentative moves. To mention but one, passages of *Human Sexual Response* are contrasted with selected passages due to heroes of past science such as Galilei and Newton. The juxtaposition aims at creating the impression that today's scientific experts write in a dry, impersonal, and dull fashion, while great scientists of the past were lively, non-technical, in-your-face writers capable of connecting with the lay public. Feyerabend takes such a difference in writing style to be a symptom of a more basic difference, that between the narrow-mindedness characterizing contemporary experts and the open-mindedness of great scientists of the past. By virtue of their being well-rounded individuals, Feyerabend maintains, the likes of Galilei and Newton are best described not as experts, but rather as dilettantes. And according to Feyerabend, "science was advanced, and is still advanced by *dilettantes*" while "experts are liable to bring it to a standstill" (1970/1999: 112). However, as Preston argues, Feyerabend can only reach his conclusions concerning the different writing styles of experts and dilettantes at the price of ignoring the context in which *Human Sexual Response* appeared, its intended audience, and the liberating effect that it exerted on the way in which sex is discussed in public discourse in the US. Moreover, the impression that Galilei and Newton's writings are easily accessible to the layperson arises because Feyerabend selectively quotes from the non-technical parts of their technical works. And of course, the claim that Galilei and Newton were not experts in their field is utterly untenable.

In terms of critical engagement with Feyerabend's views on experts, Preston's paper offers much more than the short summary above may suggest. Here we would like to draw the attention to one of Preston's take home messages. As Preston hypothesizes, it may well be that when Feyerabend railed against contemporary scientific experts, he had in mind a particular kind of expert, namely, the prototypical (CIA-funded) technocrat to whom at least some of Feyerabend's critical remarks look easily applicable. However, as Preston analysis shows, it is impossible to take the bulk of Feyerabend's critique of experts and simply apply it in our current context.

The question then naturally arises of whether one can incorporate at least some elements of Feyerabend's critique into a constructive philosophy of scientific expertise. In "Wisdom, Scientific Expertise, and Laypeople. Some Remarks on Paul Feyerabend's Philosophy of Expertise", Pierluigi Barrotta and Roberto Gronda answer the question in the positive. Scientific experts

engaged in policy advice, Barrotta and Gronda emphasize, cannot simply apply existing theoretical knowledge to particular public problems. Quite on the contrary, in order to contribute to the solution of such problems, scientific experts need to acquire the new knowledge that allows them to properly tackle the – often unique – features of the concrete contexts in which they operate. In light of the social values at stake in each specific situation, good scientific experts must be able to help in such crucial tasks as, among others, specifying values, justifying the ends pursued by a certain policy, and handling conflicts between potentially conflicting ends. Thus, good scientific experts cannot take refuge in the specialized knowledge of their discipline. Rather, they need to exhibit a feature that is often overlooked in the description of a good scientific expert. As Barrotta and Gronda put the point, they need to be endowed with wisdom – not unlike the well-rounded individual that Feyerabend envisioned as playing a central role in a free society.

The philosophy of expertise defended in Barrotta and Gronda's essay includes also a second component that has a distinctively Feyerabendian flavor. In a fully-functioning democratic society, laypeople have an epistemic contribution to make to the solution of public problems. In fact, the very definition, let alone the possible solution of a public problem hinges on how one defines the relevant concepts. And in many interesting cases, the concepts involved are "thick". For instance, such concepts as biodiversity, risk, and sustainability are thick scientific concepts, such that for the purposes of the public problems in which they are of central importance, their meaning is only partially fixed by scientific inquiry proper. In order to fully fix their meaning, one needs to adopt a normative standpoint, and take into account how lay citizens use the concepts in the understanding of their own actions and social interactions. This means that, as classical pragmatism teaches, when public problems are at stake, the community of inquirers is wider than that of scientific experts: lay citizens have the right to have a say on the meaning of the concepts, in the concrete context in which they are applied. As Barrotta and Gronda argue, the very notion of a community of inquirers that includes as its full members not only scientific experts, but also citizens, politicians, stakeholders, etc. promotes a form of pluralism that can perform the function of a shield against the temptation of scientists to gain complete control over the public sphere – just like in Feyerabend's model of a free society.

The inextricable connection of the epistemic and the normative dimensions of science-based policy advice on which Barrotta and Gronda draw the attention also lies at the heart of Karim Bschir's "Feyerabend's Humanitarian Pluralism and Its Relevance for Science-Based Policy".

Bschir provides the reader with an in-depth examination of an issue that

has so far remained largely unexplored in Feyerabend scholarship, namely, the basis for the claim repeatedly made by Feyerabend that a pluralistic method encouraging variety of opinion is the only one "*compatible with a humanitarian outlook*" (1963/1999: 97). As Bschir documents, since the early 1960s Feyerabend articulated his defense of pluralism via a sustained critical engagement with a number of positions within the philosophy of science – notably Kuhn's theory of scientific change – to which he referred as forms of "monism" and which he scolded for encouraging dogmatism, indoctrination, tyranny, authoritarianism, narrow-mindedness, and anti-humanitarianism. In his battle against monism, Feyerabend then deployed epistemological arguments, which however he supplemented, from the very start, with genuinely normative considerations: Feyerabend unequivocally claimed that pluralism is superior to monism also from the political and social point of view, and that any monistic position should be rejected on normative grounds. As Bschir shows, Feyerabend's defense of pluralism was driven by the assumption that the adoption of a certain set of rules for scientific inquiry influences both the kind and the content of the knowledge that inquiry yields. And since that knowledge has implications for society, the decisions made while choosing a certain methodology instead of another must be assessed on ethical grounds.

While consistently harsh in denouncing the flaws of monistic accounts of science, Feyerabend never addressed in a systematic way the question of how precisely one should understand the self-ascribed humanitarianism of his pluralism. Bschir's discussion of his writings nevertheless unearths certain features that reoccur in Feyerabend's description of the beneficial features of pluralism, such as, among others, the emphasis on the cultivation of individuality and the free development of individuals and the tolerance of heterodox, minority views. Moreover, Bschir argues, despite the absence of a full characterization of the humanitarian nature of his pluralism on the part of Feyerabend, his works defend a number of normative principles that are of great value for anyone interested in putting forward an account of science-based policy advice. For instance, such Feyerabendian themes as fallibilism, theory proliferation, and the extension of pluralism beyond the realm of scientific inquiry provide one with the building blocks of the framework required to properly tackle issues of urgent practical interest, such as the response of governments to the COVID-19 pandemic.

Luca Tambolo and Gustavo Cevolani's article also concerns the concrete applicability of Feyerabend's views on pluralism. Their "Feyerabendian Pluralism in Practice: Lessons from the Di Bella Case" is a deep dive into the story of a previously unknown Italian physician who in 1997-1998 became a national celebrity thanks to his alleged ability to treat cancer with an unconventional method of his own devising – the so-called "Di Bella Method"

(DBM) – presented as a radical alternative to the current standards of care, especially chemotherapy.

Indeed, as Tambolo and Cevolani detail, the DBM exhibited some of the features that Feyerabendian pluralists are bound to find attractive in an alternative to a dominant theory. And in light of Feyerabend's insistence that in order to severely test any theory *T*, one should deploy at least one alternative to *T*, it is easy to imagine a Feyerabendian pluralist who, faced with the Di Bella case, agrees with the decision made by the Italian Parliament to authorize a series of state-funded phase II clinical trials of the DBM. The decision, however, was to say the least troublesome from the ethical, scientific, and economic point of view. In fact, there was no publicly available evidence of the alleged efficacy of the DBM. Moreover, in the view of the relevant community of experts, there was no reason whatsoever to presume the DBM's possible efficacy. And as it later turned out, Di Bella's claims concerning the several thousands of patients who had undergone successful treatment with the DBM were unsubstantiated. The right thing to do in 1997-1998, Tambolo and Cevolani argue, would have been to vote against the authorization of the trials. But Feyerabendian pluralists could have recommended the correct decision only at the price of dropping the view of knowledge as an ever increasing ocean of mutually incompatible alternatives defended by Feyerabend since the mid-1960s. Tambolo and Cevolani then conclude that in order for pluralism to deliver the goods that Feyerabend claimed it can help us to achieve, Feyerabendian pluralists need to do better than Feyerabend himself did, and stick to the more limited, moderate version of pluralism espoused in some of his seminal papers of the early 1960s.

Overall, the papers included in the present Focus suggest that a satisfactory account of expert policy advice requires refined and (sometimes substantially) ameliorated versions of some Feyerabendian insights on the importance of pluralism and on the proper role of scientific experts in public decision-making. This is precisely what Piero Avitabile and Alessandro Demichelis set out to offer in their "Expanding Epistemic Public Trust. What Role for Expert-Lay Communication?". Taking their cue from the distinction between use of experts, trust of experts, and reliance on experts drawn by Feyerabend, Avitabile and Demichelis first introduce the notion of epistemic trust and then go on to discuss some of the difficulties that arise within the relationship between experts and laypeople when the passage of information is involved. As they illustrate, disagreement among experts is a particularly troublesome scenario for laypeople, who by definition lack the competence required to decide for themselves on the issues on which experts disagree. Laypeople then need to resort to "second order" evidence, that is, external indicators and sociological

proxies that allow one, if fallibly, to identify actual expertise, and therefore to assign epistemic trust.

As a blooming literature attests, the best way to systematize the conditions that should be met in order to promote the public's trust in experts is the subject of quite some controversy. Avitabile and Demichelis critically assess the proposal put forward by Gürol Irzik and Faik Kurtulmus (2019), which they claim is largely correct and yet incomplete, since it does not fully take into account epistemic responsibility. Epistemic responsibility, Avitabile and Demichelis argue, is a crucial dimension of an expert's trustworthiness, one which manifests itself in argumentative moves made by the expert operating in the public arena – moves that laypeople can directly assess. In particular, an epistemic responsible expert resorts to argumentative moves that do not hamper the public discussion and expose the expert to the risk of being contradicted not only by other experts providing contradictory testimony, but also by lay citizens, who can report the testimony of other experts. More specifically, on Avitabile and Demichelis' account, what is required of an epistemically responsible expert is the providing not of categorical testimony of the form "p", but rather, of reasoned testimony of the form "p, because q", which allows the discussion to move forward, for instance by the posing of questions concerning q. This way of operationalizing the notion of epistemic responsibility, Avitabile and Demichelis maintain, allows one to properly acknowledge laypeople's role as legitimate interlocutors in public debates – a key ingredient of Feyerabend's free society.

As Avitabile and Demichelis openly declare, they are much more confident than Feyerabend ever seemed to be in the possibility to build actual, substantial public trust in scientific experts. Their essay then exemplifies a thread that runs through all the contributions to the present Focus. In fact, the papers briefly introduced here vividly show that despite their flaws and limitations, Feyerabend's ideas continue to be relevant to this day, since they provide us with precious material to continue to perfect our thinking on, among many other things, the role of experts in democratic societies.

Roberto Gronda
University of Pisa
roberto.gronda@unipi.it

Pierluigi Barrotta
University of Pisa
pierluigi.barrotta@unipi.it

Luca Tambolo
Independent scholar
ltambolo@gmail.com

Gustavo Cevolani
IMT School for Advanced Studies Lucca
gustavo.cevolani@imtlucca.it

References

Feyerabend, Paul Karl, 1963/1999, "How to Be a Good Empiricist: A Plea for Tolerance in Matters Epistemological", as reprinted in Feyerabend 1999: 78-103.

—, 1970/1999, "Experts in a Free Society", as reprinted in Feyerabend 1999: 112-126.

—, 1978, *Science in a Free Society*, New Left Books, London.

—, 1999, *Knowledge, Science and Relativism: Philosophical Papers, Volume 3*, edited by John Preston, Cambridge University Press, Cambridge.

Irzik, Gürol, Faik Kurtulmus, 2019, "What Is Epistemic Public Trust in Science?", in *The British Journal for the Philosophy of Science*, 70: 1145-1156.

Masters, William H., Virginia E. Johnson, 1966, *Human Sexual Response*, Little, Brown and Company, Boston.

Feyerabend, experts, and dilettantes

John Preston

Abstract: Paul Feyerabend's 1970 article "Experts in a Free Society" tries to make the case that scientific experts can only be tolerated if they are *dilettantes*. He uses Galileo, Newton and Kepler as examples of great scientists whose writing is nothing like that of contemporary "experts", these latter being represented by the authors of the well-known book *Human Sexual Response*, Bill Masters and Virginia Johnson. He goes on to argue against the idea that the Scientific Revolution represented the triumph of empiricism.

I take issue with the way Feyerabend represents Galileo as implacably opposed to empiricism, with his supposition that good science requires a particular personality, and with the way in which he represents the work of Masters and Johnson.

Keywords: Feyerabend, Experts, Dilettantes and amateurs, Empiricism, Scientific Revolution.

Introduction

Paul Feyerabend is now best known for his book *Against Method* (1975), in which he tried to alter our perception of science and the lessons we should learn from it. On the basis of an historical case-study of Galileo, he argued that there is no such thing as scientific method, since great scientists are methodological opportunists who use any strategies they can in order to promote their preferred theories. He didn't talk much about scientific experts in that book, but in other work he published during that decade he had quite a lot to say about them, and what he said bears even more investigation nowadays, when we rely on experts for their views on climate-change, COVID-19, vaccination, etc.

Most of what Feyerabend said about experts fell somewhere between the sceptical and the scathing. So, for example, he took delight in pointing out that "obstinate and conceited" experts can be outsmarted by laypeople (notably, by lawyers), that experts often disagree with one another, that their judgements can legitimately be challenged by non-experts, that they can legitimately be overruled by political interference, and he was happy to refer to them as "know-it-alls".

Feyerabend also rarely said anything positive about the people he thought of as scientific experts. However, going somewhat against his popular reputa-

philinq XII, 1-2024, pp. 89-106
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.553
Submitted: 7 April 2024
Accepted: 27 June 2024

tion, he was capable of saying very positive things about science. How could this be? An answer lies in the paper I will discuss.

The context of Feyerabend's career

By the end of the 1960s, Paul Feyerabend's star had risen very high. He had made his mark with important papers published in the late 1950s, while he was still associated with the Popperian school, defending scientific realist approaches to scientific theory against positivist approaches. But in the early 1960s he had broken with Popper, publishing papers which were critical of the hypothetico-deductivist approach to methodology which was common to Popperians and to logical empiricists. Together with Thomas Kuhn, he had also been at the forefront of a new *historical* approach to philosophy of science, thus moving away from the logical approaches that were, again, common to Popper and the logical empiricists. His papers on Niels Bohr had been critical of Popper, and although he wasn't able to present his paper to the International Colloquium in the Philosophy of Science at Bedford College, London, in 1965,[1] he did present it to Popper's LSE seminar in 1967, and when it was published in one of the resulting volumes, "Consolations for the Specialist" (published in *Criticism and the Growth of Knowledge*, 1970) was very critical of Popper, *and* Kuhn, *and* Imre Lakatos. At the end of that paper, Feyerabend concluded that no proposed method for rationalizing science can be applied, that "what remains are aesthetic judgements, judgements of taste, and our own subjective wishes" (1970b/1981: 160), and further that "the choice between theories which are sufficiently general to provide us with a comprehensive world view […] may become a matter of taste" (1970b/1981: 160). In that same year, in the original, essay version of *Against Method* (1970d) Feyerabend had argued that there is no scientific method.

He expanded on this in a shorter paper entitled "Experts in a Free Society" (1970c/1999), and brought it to bear on the issue of expert scientific authority, which clearly bothered him. What follows is a critical examination of some themes from this paper, although I'm not going to cover those parts of it that make Feyerabend's case "against method", since that has been covered extensively elsewhere.

"Experts in a Free Society" (1970)

"Experts in a Free Society" was published, not in an academic journal, but in a magazine. The magazine in question, *The Critic*, was produced by the

[1] For the full story, see Collodel *et al.* (2020: 402n).

Thomas More Association, a non-profit organisation based in Illinois. It's hard to find any information about this organisation or its magazine online, since there now exists a well-known (and very web-assertive) Thomas More *Society*, a conservative Roman Catholic public-interest law firm based in Chicago, which was founded much later. In 1962, though, the *Catholic Standard and Times* reported that "A magazine founded and edited by laymen has won for the second consecutive year the Catholic Press Association's general excellence award in competition among general interest magazines. The award went to *The Critic*, self-described as a 'magazine of Christian culture.' It is published six times a year by the Thomas More Association of Chicago" (The Catholic News Archive, online). *The Critic* was published from 1948 to 1981, and then from 1985 to 1990.

Feyerabend prefaced this article with a slogan ("Jargon, jargon everywhere. And all of it doth stink"), a warning (that the paper will sound a little harsh, and unjust), and a "confession" (no doubt with the context of publication in mind) explaining that it was written "in a fit of anger and self-righteousness caused by what I thought were certain disastrous developments in the sciences". The key lies in his view that experts are the problem with science, rather than its lifeblood (or even a desirable feature of it at all):

> I think very highly of science, but I think very little of experts, although experts form about 95% or more of science today. It is my belief that science was advanced, and is still being advanced by *dilettantes* and that experts are liable to bring it to a standstill (1970c/1999: 112).

This I take to be the *main* thesis of his paper. I'll give a brief exposition of some themes from "Experts in a Free Society", and then address some of the issues they raise.

Feyerabend's attitude to experts

Feyerabend begins, before explaining what experts *are*, by setting out his feelings about them. If, he argues, our attitude towards giving people the vote is as strict as it is today, then we should *not* give experts the vote, because of their immaturity. (In the US at that point, there was public discussion of whether eighteen-year-olds should be able to vote. The Twenty-sixth Amendment, lowering the voting age from twenty-one to eighteen, was ratified only in 1971). Rather, we must wait until they grow up, become mature and responsible, and this means becoming what Feyerabend calls a *dilettante*. (We will soon investigate what he means by this term, but for the moment we can understand it as meaning a non-specialist). In a *free* society, though (I presume that this is the contrast he's mak-

ing) experts will certainly have the vote; they will certainly be listened to as every other citizen will be listened to, but they will receive none of the special powers which they would so dearly love to possess. Laymen will look after their affairs and will make the decisions which must be made if we want to apply science to society. Laymen will control science – and no harm will come of it.

What, then, *are* experts? Experts are people who have decided to achieve excellence in a narrow field "at the expense of a balanced development" (1970c/1999: 113). They subject themselves to standards which restrict them in many ways, conduct their working lives in accordance with these standards, and divide their lives into separate domains. This has bad consequences: special subjects are voided of ingredients which make a human life beautiful and worth living; these ingredients are impoverished, too; emotions become crude and thoughtless, just as thoughts become cold and inhumane. Professionalism takes over.

Feyerabend gives us examples of this professionalism, but prefaces these by discussing some examples of science writing that *don't* exhibit it. His examples are drawn from Galileo, Kepler, and Newton. I will focus on Galileo and Newton, the two figures most important for what follows (as well as for *Against Method*, in Galileo's case, of course).

Galileo's acquaintance with the telescope

Feyerabend quotes two passages from *Sidereus Nuncius* (*The Starry Messenger*), written in 1610, the longer one being:

> About ten months ago a report reached my ears that a certain Dutchman had constructed a spyglass by means of which visible objects, though very distant from the eye of the observer, were distinctly seen as if nearby. Of this truly remarkable effect several experiences were related, to which some persons gave credence while others denied them. A few days later the report was confirmed to me in a letter from a noble Frenchman at Paris, Jacques Badovere, which caused me to apply myself wholeheartedly to inquire into the means by which I might arrive at the invention of a similar instrument (Drake 1957, ed.: 28-29).

What we have here, Feyerabend explains, is "a personal story, a very charming personal story which slowly leads us to the discoveries, and these are reported in the same clear, concrete, colourful way" (1970c/1999: 114).

Newton's experimental preparation

Feyerabend is also impressed by a passage from the beginning of Isaac Newton's first paper on colours in which Newton (who, as we shall see, Feyerabend

usually has no time for), describes procuring a glass prism, preparing his chamber for the experiment, and viewing "the vivid and intense colours produced" when a beam of light hits the prism and is refracted onto the opposite wall:

> [I]n the beginning of the year of 1666 […] I procured me a triangular glass prism, to try therewith the celebrated phenomena of colours. And in order thereto having darkened my chamber and made a small hole in my window shuts, to let in a convenient quantity of the sun's light, I placed my prism at his entrance, that it might be thereby refracted to the opposite wall. It was at first a pleasing divertissement to view the vivid and intense colours produced thereby; but after a while applying myself to consider them more circumspectly, I became surprised to see them in an *oblong* form… ("The New Theory about Light and Colours", 1672, as reprinted in Thayer, 1953, ed.: 68).

Feyerabend sums up his three examples by asking us to remember that

> all these reports are about cold, objective, "inhuman", *inanimate* nature; they are about stars, prisms, lenses, the moon; and yet they are described in a most lively and fascinating manner, communicating to the reader an interest and an excitement which the discoverer felt when first venturing into strange new worlds (1970c/1999: 115).

Professional science writing: Masters and Johnson

We then get Feyerabend's two examples of *bad* science writing, or at least bad writing: the well-known book *Human Sexual Response*, by William Masters and Virginia Johnson, and a pamphlet containing the report of an *ad hoc* committee examining allegations of police brutality on the Berkeley campus during 1968-1969. The latter would have been of personal interest to Feyerabend, of course, since he had a role in the student unrest. But I can't find any copy of it (online or elsewhere), so will discuss the former.

Human Sexual Response was published in 1966, and had become something of a *cause célèbre*. Since the book is, as Feyerabend puts it, "about men rather than about stones or prisms", one would expect a beginning even more lively and interesting than that of Galileo, Kepler, or Newton. But if we turn to it we find things like this…

> "In view of the pervicacious gonadal urge in human beings, it is not a little curious that science develops its sole timidity about the pivotal point of the physiology of sex. Perhaps this avoidance…" and so on. This is no longer human speech. This is the language of the expert (1970c/1999: 115).[2]

[2] The horrible English word "pervicacious" means obstinate or stubborn. The "gonads" are the primary reproductive organs: the testes in the male, and the ovaries in the female.

I won't regale you with the other things Feyerabend quotes, but simply relate his observations. First, the *subject*, the human person, has left the picture. Second, relatively simple human matters are put in a more wordy and more complicated way than those matters demand. Third, "irrelevant technical terms intrude and fill the sentences with antediluvian barks, grunts, squeaks, belches" (1970c/1999: 115). The effect is that

> [a] wall is erected between the writer and his readers not because of some lack of knowledge, not because one does not know who the reader is, but in order to make utterances conform to some curious professional ideal of objectivity. And this ugly and inarticulate idiom turns up everywhere and takes over the function of the most simple and straightforward description (1970c/1999: 115).

Feyerabend's complaints don't stop there. Instead of saying that when masturbating, a woman will stop only when she gets tired, Masters and Johnson say: "usually physical exhaustion alone terminates such an active masturbatory session" (1966: 65). And instead of saying that the man should *ask* the woman (whom Feyerabend himself calls "the female") what she wants or does not want rather than try to guess it on his own, Masters and Johnson say: "The male will be infinitely more effective if he encourages vocalisation on her part" (1966: 66). We shouldn't pretend that this is a matter of their wanting to be more precise (their expression "*infinitely* more effective" precludes that), and neither should we excuse it by claiming that they want to address their fellow professionals rather than the general public and so have to use a special lingo in order to make themselves understood: "[W]e are not dealing with the structure of organs, or with special physiological processes which might have a special name in medicine, but with an ordinary affair such as *asking*" (1970c/1999: 116). Feyerabend concludes:

> Galileo and Newton could do without a special lingo although the physics of their time was highly specialised and contained many technical terms. They could do without a special lingo because they wanted to start afresh. Masters and Johnson find themselves in the same position, but they cannot speak straight any more, their linguistic talents and sensibilities have been distorted to such an extent that one asks oneself whether they will ever be able to write normal English again (1970c/1999: 116).

And after his discussion of the Berkeley pamphlet Feyerabend concludes that in these both texts:

> a single and rather impoverished idiom takes over all functions and is used under all circumstances. Does one want to insist that the thought that hides behind this ugly and inhuman exterior (emphasis upon abstract processes such as "communication" instead of living people) has remained nimble and humane? (1970c/1999: 117).

Free men (dilettantes) vs. *slaves, and the power of experts*

According to Feyerabend, it was already realised by Aristotle that being an expert was a *predicament*, and not a matter of pride (1970c/1999: 117). He summarises Aristotle's picture of "the free man" as a man having a sense of balance and perspective, being well-informed in different aspects of human culture, having emotions, and having concrete experience of what goes on outside the limits of his own subject. And Feyerabend's way of putting this is: you can be a free man, you can achieve and yet retain the dignity, the appearance, the speech of a free man only if you are a *dilettante* (1970c/1999: 117-118).

Feyerabend asks us to remember "to what extent the academic profession makes slaves of its members" (1970c/1999: 118). He finds support for this in the sociologist Robert Merton, who had recently characterised science as "a system of institutionalised vigilance, involving competitive co-operation" (1969: 220). Such a tough-minded or "realistic" picture of contemporary science, then, means that "wandering minstrels who try to bewitch the onlookers by praising the beauties of science, the joys of discovery, the essentially human character of the search for knowledge and truth […] are singing about a time that has long gone by [i.e., the heroic time of the 'Scientific Revolution']" (1970c/1999: 118). (I imagine that Feyerabend had all sorts of people in mind here, including Popper and Lakatos).

Feyerabend sums up: experts today are excellent, useful, irreplaceable, but mostly nasty, competitive, ungenerous slaves, slaves both in mentality, speech, and in social position (1970c/1999: 118). But these inarticulate and slavish expert minds now have power over large parts of society. What arguments lie behind this claim for experts to have such power? (1970c/1999: 119). Feyerabend proposes to investigate, and he does so by pivoting to discuss the character of the Scientific Revolution.

The Scientific Revolution, Galileo, and empiricism

Many people, he claims, still believe that the Scientific Revolution of the 16th and 17th centuries was the result of a "radical" empiricism and think that it occurred only because one man decided to eliminate views not in agreement with observation and reasonable generalisations therefrom. But the Scientific Revolution *wasn't* the result of a radical empiricism. Galileo was suspicious of experience, since experience played a role in the magical tradition which he despised. Alchemy, too, and the art of discovering witches, were firmly empirical. Galileo was suspicious about "[s]uch magical and enthusiastic types of empiricism which exclude thought from a large area of knowledge" (1970c/1999:

120). He even rejected the more sober empiricism of Aristotle (which Feyerabend takes to be the most promising kind).

Feyerabend agrees that the Scientific Revolution was the heroic time of science, but he insists this is because the best scientists were the best *dilettantes*. In this scenario, expert knowledge is produced, but it is produced *by* dilettantes. By "dilettantes", Feyerabend seems to have meant *non*-specialists, thinkers who did not limit their thought to any single area.

Evaluation of Feyerabend's attitudes and claims

How, then, should we evaluate these attitudes and claims on view in "Experts in a Free Society"?

Feyerabend wasn't really interested in explaining what experts *are*. One might sympathise with his admiration for "Renaissance men", i.e., polymaths (and I do mean *men* – women hardly figure in Feyerabend's works). But narrowness of excellence will make "experts" of almost every significant contributor to any intellectual field (including Feyerabend himself). At that point one has to ask: do the consequences that Feyerabend supposed really follow? Perhaps Feyerabend's critique is best thought of as a blast at a certain *kind* of expert prominent in the cold war US context: the (CIA-funded) technocrat, keen on presenting humans as machines. That kind of expert is less fashionable or perhaps just more sophisticated nowadays, even in the US, and so one might wonder whether experts *have* grown up, become mature and responsible. Certainly I think one can't just take Feyerabend's case against them and assume that it applies with full force to our current context.

Feyerabend's (early) Galileo

It's true that what Feyerabend quotes from *Sidereus Nuncius* is "a personal story, a very charming personal story which slowly leads us to the discoveries, and these are reported in the same clear, concrete, colourful way" (1970c/1999: 114). But Galileo was in a very particular and unusual context: first, his text is dedicated to a particular person, his patron, the Grand Duke of Tuscany. That demanded something of a personal manner. Second, he obviously felt the need to explain to his reader (and his patron) how he came to know about this remarkable and controversial new instrument, and how one might construct a similar instrument. Patronage wasn't untypical, of course, but the *combined* context is untypical enough that it's unfair to demand that personal stories should accompany or preface all scientific communications, or even to suppose that they could do so.

What about the way in which Galileo writes? Lively and personal writing impresses Feyerabend, of course, and he himself almost always wrote in that way, but it seems to be a pretty superficial criterion for scientific merit. It's true that Galileo and Newton could do without a special lingo in *preambles to* their physical works, but the physics of their time was highly specialised and contained many technical terms. The extent to which they wanted to "start afresh", as Feyerabend supposes, is contestable. They didn't eschew the existing technical terms in mechanics, optics, etc. Feyerabend is impressed by scientists who write in ordinary language. And surely there's lots of bad writing in science texts. But it would be wrong to issue a general prohibition on jargon (i.e., technical terms) in science. Feyerabend's anti-jargon slogan would apply to *all* science, since all sciences (and all scientists, including the ones Feyerabend revered) use *some* technical terms.

Galileo against empiricism?

What about Feyerabend's insistence that Galileo rejected empiricism? There's clearly *something* right about this, but also, something clearly wrong.

No-one reading *Sidereus Nuncius* today could fail to think of it as empiricist in a *broad* sense. Its whole *raison d'être* is to relate several series of observations of heavenly bodies that Galileo claims to have made with his "spyglass" (telescope). So what can Feyerabend have meant (here) when he claimed that Galileo wasn't an empiricist?

To see this, we have to turn to the contexts in which each of them was writing. *Galileo's* context was characterised partly by the empiricism of the Aristotelian tradition but also (Feyerabend thinks) by a kind of "empiricism" associated with magic, alchemy, and witchcraft.

Feyerabend's context was of course one in which forms of empiricism had dominated Anglo-American philosophy generally, and philosophy of science in particular. But the forms against which Feyerabend was reacting were drawn both from recent philosophy and from past philosophies of science. One of the latter was *Newton's* empiricism, which Feyerabend railed against in an article published in the same year as "Experts in a Free Society", namely, "Classical Empiricism" (Feyerabend 1970a/1981). Among the others were the kinds of sense-datum empiricism popular among some of the founders of analytic philosophy (Russell, Moore), *then* among the logical positivists, and *then* among figures who tried to combine these approaches (such as A.J. Ayer). No doubt some of these could be characterised as "radical" empiricisms.

When Feyerabend insisted that Galileo rejected empiricism, he seems to have had both "radical" and "naïve" empiricism in mind. Feyerabend is on

strong ground when understood as claiming that Galileo wasn't a "*radical* empiricist". His methodology doesn't seem to have featured the kind of restrictive "rules" that characterised *Newton's* empiricism (and obviously he wasn't an empiricist as conceived by the tradition of British empiricist *philosophy*, or a sense-datum empiricist). But the idea that the Scientific Revolution was the result of *such* an empiricism is far-fetched, and also difficult to pin on anyone. When Feyerabend speaks of the myth that one man initiated the Scientific Revolution by eliminating views not in agreement with observation he *may* have been thinking of Stillman Drake, whose edition of Galileo's essays he was working from, and who did present Galileo as an empiricist (see, e.g., Drake 1957, ed.: 8, 18). But "*radical*" empiricism is a red herring.

So is "naïve" empiricism, assuming that this is the view that we have to take the unaided deliverances of sense-experience as the only court of scientific appeal. Again I suspect that *no-one* in science or its philosophy would have endorsed an empiricism of *that* kind (not even figures such as Francis Bacon and Ernst Mach, who are sometimes thought of in this way).

When we come to Feyerabend's claim that Galileo rejected the empiricisms of Aristotle and of the magicians, matters are more complicated.

As I just said, Galileo would have rejected the idea that *unaided* human sense-experience provides the sole evidence-basis of science. He did not, however, *reject* the appeals that Aristotelians made to sense-experience. Rather, he answered them. And, as Maurice Finocchiaro has recently pointed out, in a certain respect Galileo used Aristotle's own empiricism against Aristotle, in arguing that "it is more in accordance with Aristotle to philosophize by saying 'the heavens are changeable because so the senses show me' than if you say 'the heavens are unchangeable because theorizing so persuaded Aristotle'" (Finocchiaro 2019: 129-130, quoting Galileo 1997: 104).[3]

The Witchfinder General's "experience"

What about the magicians, alchemists and witch-hunters to whom Feyerabend adverted, and who did appeal to "experience" (and "experiment")? Was it the kind of experience to which Galileo was appealing, that is, human sense experience (as aided by his new instrument, the telescope)? Let's see. Which of Feyerabend's examples should we consider? The one that's most fun, of course.

In *The Discovery of Witches* (1647), credited to Witchfinder General Matthew Hopkins, the question is put: "From whence proceeded [the Witchfinder's] skill? Was it from his profound learning, or from much reading of

[3] For a more balanced view of whether Galileo was an empiricist see, e.g., Hesse (1967).

learned authors concerning that subject?". And the answer given there, which Feyerabend quotes, is: "From neither or both, *but from experience*, which though it be meanly esteemed of, is yet the surest and safest way to judge by" (1970c/1999: 120, Feyerabend's emphasis). But this clearly refers to what we might call "hands-on" or "lived" experience, not *perceptual* experience. And this is amply confirmed when we read on, since the questioner asks "Where was this experience gained? and why was it gained by him and not by others?", to which the reply begins "The Discoverer never travelled far for it, but in March 1644 he had some seven or eight of that horrible sect of Witches living in the Towne where he lived, a Towne in Essex called Maningtree, with divers other adjacent Witches of other towns…". (The text's four other mentions of "experience" also clearly refer to lived experience). In claiming that the discovery of witches was "firmly empirical", insinuating that *philosophical* empiricists would have to acquiesce in it, and supposing that this was the same kind of experience to which Galileo was appealing, Feyerabend is cheating.

Appeals to sense-experience in Sidereus Nuncius

There's plenty of evidence of Galileo's own (more sophisticated) empiricism (or empirical stance, we might say) in *Sidereus Nuncius*. That text begins:

> Great indeed are the things which in this brief treatise I propose for observation and consideration by all students of nature. I say great, because of the excellence of the subject itself, the entirely unexpected and novel character of these things, and finally *because of the instrument by means of which they have been revealed to our senses.* Surely it is a great thing to increase the numerous host of fixed stars previously visible to the unaided vision (Drake 1957, ed.: 27, emphasis added).
>
> In this way one may learn *with all the certainty of sense evidence* that the moon is not robed in a smooth and polished surface, but is in fact rough and uneven… (Drake 1957, ed.: 28, emphasis added).
>
> Again, it seems to me a matter of no small importance to have ended the dispute about the Milky Way *by making its nature manifest to the very senses as well as to the intellect* (Drake 1957, ed.: 28, emphasis added).
>
> All these facts were discovered and observed by me not many days ago with the aid of a spyglass which I devised (Drake 1957, ed.: 28).
>
> I have observed the nature and the material of the Milky Way. With the aid of the spyglass this has been scrutinized so directly and with such ocular certainty that all

the disputes which have vexed philosophers through so many ages have been resolved, and we are at last free from wordy debates about it (Drake 1957, ed.: 49).[4]

Of course, what Feyerabend was opposing *was* an empiricist myth, a myth according to which the Scientific Revolution occurred simply because people started observing rather than speculating. (Perhaps that's what he meant by calling it "*radical*" empiricism). No doubt this myth featured in the popular consciousness, and perhaps also in certain popularisations of the history of science. But if Feyerabend thought that it had been promulgated by serious works of philosophy, or history of science, he should at least have told us which ones.

Feyerabend and the Masters of Sex

Feyerabend's take on Masters and Johnson is equally bizarre. He misrepresents both what they write, and why they wrote the way they did. Thomas Maier's book *Masters of Sex*, which was subsequently turned into a TV series of the same name, supplies ample evidence that Masters and Johnson were not at all the kind of people Feyerabend imagined. His objections to their work completely fail to recognise its importance, cultural as well as scientific.

What Feyerabend quotes from *Human Sexual Response* (Masters *et al*. 1966: V – the passage I quoted above) is a strange way to begin the *preface* to a book, it's true. Convention demands that prefaces be reader-friendly. But what he quotes are not the words of Masters and Johnson, but of Robert L. Dickinson and Henry H. Pierson, from their article "The Average Sex Life of American Women", published the *Journal of the American Medical Association* back in 1925 (what Feyerabend quotes is from page 1113).

One might well wonder whether Masters and Johnson themselves wrote like Dickinson and Pierson. Well, despite their book having gone on to be, as Feyerabend himself notes, a best-seller (1970c/1999: 115; Maier 2009: 174, 295, 360), it's true that they did so. Although they kept their distance from the "dry clinical language" of Alfred Kinsey's prior research (Maier 2009: 97), their book was "designed primarily for doctors and residents in need of basic training" (Maier 2009: 151), and published as a medical text. Masters was more inclined to jargon than Johnson (Maier 2009: 106), and wrote in what Maier calls an "obstipated medicalese" (2009: 123). Consequently, their book is full of medical terms ("ovulatory-mucus plug", "ovulatory-type mucoid material", "cervical os", "patulous", "nulligravid cervix", "the parous individual" – all from Masters *et al.* 1966: 115), to the extent that it required a glossary (Maier 2009: 158).

[4] Note that two of these passages credit us with *sense-certainty* (although this isn't the philosophers' sense of "certainty").

However, despite his acquaintance with their context, Feyerabend fails to register their need to put their case in "language as 'non-inflammatory' as possible" (Maier 2009: 157). In the USA at that time, Masters and Johnson just could not have written about sex in ordinary and non-technical terms and had their case taken seriously, or indeed as anything other than pornographic. As Maier says:

> *Human Sexual Response* transformed the public discourse about sex in America, opening a new era of candidness never seen before in the media. Although ridiculed for their turgid prose, Masters and Johnson relied on medical terms and clinical descriptions that didn't offend readers. They stayed away from vulgar phrases that would have invited censorship (2009: 174).

Or, as Masters himself put it: "You must remember that in publishing this book we were concerned primarily with acceptance – that is the reason that it wasn't in English to start with" (Maier 2009: 174).

As for his comments on their character (their thinking not being "nimble and humane"), Feyerabend clearly knew nothing about Bill Masters and Virginia Johnson.[5] He can't be faulted for not knowing about their sexual relationship (Maier 2009: Ch. 14 onwards). But he didn't know that neither of them was trained in psychology or psychiatry (Maier 2009: 179). I guess that, pre-internet, we didn't always have time to research the people whose books we were writing about. But let's bear in mind that *he* was happy to accuse contemporary scientists of knowing nothing about the research *they* denigrated (into astrology, Chinese medicine, etc.). He was just attacking a target that he himself had stereotyped, and using a double-standard in doing so.

Experts, dilettantes, and slaves

I have no quarrel with what Feyerabend quotes from Aristotle's discussion of the "free man" (*Politics*, Book VIII, section 2). Indeed I think it may be an important ideal, more widely applicable than Aristotle would have thought.

However, I do want to challenge Feyerabend's idea that one can be an Aristotelian "free man" only if one is a *dilettante*. This is completely wrong. Aristotle's ideal is similar to what north Americans call a "well-rounded" person, *not* a "dilettante", and Aristotle would *not* have accepted the equation of free men and dilettantes, at least not using "dilettante" in its ordinary Italian *or* English senses. In Italian it means amateur. But Galileo was not an *amateur* scientist!

[5] By the time he was writing, they were working at an independent not-for-profit research institution they had founded in St. Louis in 1964, originally called the Reproductive Biology Research Foundation.

He'd studied mathematics at the University of Pisa, and obtained professorial positions in Pisa and then Padua, before moving to Florence as Philosopher and Chief Mathematician to the Grand Duke of Tuscany. No-one, at that time, could have been less of an amateur.

In the English sense a dilettante is a person who has a *shallow* interest in some area of culture or knowledge. So the relevant contrast is shallowness vs. depth, not narrowness vs. breadth. An expert, that is, someone who has a *deep* knowledge of something, can quite well be the kind of person who also has *broad* interests and competences, i.e., a "free man" in Aristotle's sense. The two are orthogonal.

On a more positive note, perhaps what Feyerabend had in mind was that being a dilettante involves having a kind of naïveté (or openness) towards the knowledge acquired in the field, leading the researcher – however a well-known specialist they might be – to scrutinize established knowledge without taking it for granted.[6] On such an understanding of the term, a specialist would approach the subject as if he or she was a dilettante, raising questions and concerns that might seem naïve and provocative for other researchers, but that are capable of unveiling new paths of inquiry. As we will see in a moment, this is a better fit with the kind of scientists Feyerabend most admired.

As for Feyerabend's claim that the academic profession makes slaves of its members, though, here we should just recall that Aristotle thinks *all* paid employments, even at the University of California, Berkeley, "absorb and degrade the mind" (quoted in Feyerabend 1970c/1999: 118). Aristotle provides no basis for what one might *like* to see here, that is, a critique of the way some people are treated within and on the fringes of academia (some people, but *not* others: even academics who don't have Paul Feyerabend's astronomical salary are still well-treated relative to other kinds of employees).

Galileo as a dilettante?

Neither Galileo's sophisticated empiricism, that is, the combination of realism with empiricism that his view exhibits, nor his not being a naïve or radical empiricist, means that he was a dilettante in any ordinary sense, for sure.

Was he, though, a dilettante in Feyerabend's sense, that is, an Aristotelian "free man"? (I'm not trying to make a joke at the expense of the Catholic Church here). Was he a well-rounded individual? As far as we can tell from biographies, yes. And surely that helped him to *write* well. But this is only a part of what made him a great scientist, and of what made his contributions to science great ones.

[6] For this suggestion I am grateful to an anonymous referee.

Certainly the ranks of great scientists feature their share of narrowly-focussed single-minded individuals, alongside plenty of more well-rounded ones. Nothing can be said against the list of Feyerabend's most prominent scientific heroes (which would include Aristotle, Galileo, Ernst Mach, Ludwig Boltzmann, Albert Einstein, Max Born, and Niels Bohr). Or against his list of scientific villains, which would no doubt feature Isaac Newton (Feyerabend calls him "The awful Newton, who more than anyone else is responsible for the plague of professionalism from which we suffer today"), and Max Planck. (It would also have included a host of minor figures – in fact, this category included almost *all* scientists). But (leaving aside issues of nationalism, I hope) here we really *do* have a *subjective* preference, and not a good basis for evaluating whose science was good or who was a great scientist.

Rather than Feyerabend's personality monism (everyone, or at least every expert, should be an Aristotelian "free man"), I think we ought to prefer a *pluralism* of personality-types, combined with a social mechanism that ensures that no one type of personality has the sole say in public-facing matters.

The advancement of science

But there is a deeper issue: what are we to make of Feyerabend's picture of how science advances? He assures us that science was advanced, and is still being advanced by dilettantes, and that experts are liable to bring it to a standstill. But what criterion could he give for science being either "advanced" or "at a standstill"? I strongly suspect that the answer is: none! As long as there are experts for each alternative theory in the "ocean of alternatives", his conception pretty much precludes it.

His perspective on this is worth comparing with that of his friend and contemporary Thomas Kuhn. It was Kuhn who popularised the notion of scientific *revolutions* (revolutions, *plural*, that is, relatively swift and wholesale revisions of the basic assumptions of a given scientific field) in his famous book *The Structure of Scientific Revolutions* (1962).

Kuhn would insist that most science is not *and cannot possibly be* a matter of revolutionary thinking. For Kuhn, most science takes place under the auspices of a "paradigm", a set of deep assumptions about the domain in question, the sort of assumptions which science students learn when they are being educated into the field, and which cannot be the subject of constant questioning. Science *advances*, though, during these periods of "normal science", just as surely as it does in "scientific revolutions". The kind of change in question is cumulative, rather than revolutionary, but it is just as real. And this kind of science, normal science, is the home playing-field of the expert.

I suspect that Kuhn would also insist that those who bring about scientific change don't do so because they are dilettantes. For Kuhn, scientific revolutionaries are, as it were, scientific conservatives who are *at the end of their rope*. In no way are they trying to be revolutionary – his perspective would scorn efforts to deliberately "revolutionise" scientific thinking. Those who initiate scientific revolutions, for Kuhn, are scientific experts who are trying to *defend* their existing paradigm, a paradigm which by then is in a serious state of crisis, by reinterpreting or re-jigging it in various ways. If that results in a scientific revolution, that is very much the *unintended* effect of their non-revolutionary activities.

Kuhn's picture may not be right. Maybe it fits only some of those who initiated scientific revolutions (like Charles Darwin, or Planck, who Feyerabend thought of as a dogmatist), and not others (such as Einstein). In addition, though, we must ask Feyerabend: why, when it comes to expert opinion, should we think that it is only the "*advancement*" of science that matters? Even if some particular science had stagnated, a situation which would appal Feyerabend, might it still not give the best answers to questions about the domain in question? Feyerabend didn't much care about the periods of stability between scientific revolutions, and most of his scientific heroes were people who presaged or brought about such revolutions. But why, when it comes to knowledge and the expert advice that should be based upon it, should we think that only *change* counts? Might not good expert advice be based upon scientific theorising which is (at that time) stable? In fact, isn't that what we generally expect, and *want*? Revolutionary science may feature experts, but one might well hesitate to take the advice of experts in such turbulent times.

Conclusions

"Experts in a Free Society" is lots of fun, but unreliable. Feyerabend was mainly concerned with the way scientists *write* (rather than the conclusions they come to, for example), and he was *generally* good at characterising that. But his claims on behalf of Galileo and Newton, who come away well, are exaggerated: each of them *did*, once Prefaces were done with, use technical language (physicists always do). That Galileo and Newton were (in certain respects) "starting afresh" should come as no surprise – but not every scientist can be expected to do so, and it would be a disaster if they did do so.

Feyerabend did a disservice to Masters and Johnson. He failed to find out about the two people in question, and ignored the context and the (mostly positive) effects of *Human Sexual Response*. Perhaps his critique is best thought of as a blast against a certain *kind* of expert, the kind then being produced in the US context. His vitriol about experts is understandable in

that context, but not generalisable (and Masters and Johnson were not experts of *that* kind).

Feyerabend was concerned with a popular myth about the Scientific Revolution, but he failed to show that this process had nothing to do with empiricism, since the kinds of empiricism he critiqued are *too philosophical*. He didn't successfully show that Galileo would have opposed the (now commonsensical) kind of empiricism that *did* play a role in the Scientific Revolution, and he ignored the things in Galileo's text which point to him having *not* rejected empiricism. His appeal to witch-hunters, magicians, alchemists, etc. was fanciful since, even if Galileo was suspicious of them, they weren't empiricists of the relevant kind.

Aristotle's "free man" may be a worthy ideal, but it is *not* to be identified with the dilettante (or the opportunist). Aristotle would not have accepted that equation, and none of Feyerabend's hero-scientists were dilettantes. Expertise is *deep* knowledge, and that's perfectly compatible with having a *wide* range of interests, as well as with the ability to see things in fresh and unprejudiced ways.

Finally, Feyerabend was preoccupied with how science advances, and that is always a one-sided preoccupation. His claims that science was advanced, and is still being advanced, by dilettantes, and that experts are liable to bring it to a standstill are problematic, since he offers (and may be *able* to offer) no criterion for science being either "advanced" or "at a standstill". In fact, by the time of "Experts in a Free Society", he had deprived himself of any way of saying how or whether science advances.

John Preston
University of Reading
j.m.preston@reading.ac.uk

References

Collodel, Matteo, Eric Oberheim, 2020, eds., *Feyerabend's Formative Years, Volume 1: Feyerabend and Popper – Correspondence and Unpublished Papers*, Springer, Cham.

Dickinson, Robert L., Henry H. Pierson, 1925, "The Average Sex Life of American Women", in *Journal of the American Medical Association*, 85: 1113-1117.

Drake, Stillman, 1957, ed., *Discoveries and Opinions of Galileo*, Doubleday, New York.

Feyerabend, Paul Karl, 1970a/1981, "Classical Empiricism", as reprinted in Feyerabend 1981: 34-51.

—, 1970b/1981, "Consolations for the Specialist", as reprinted in Feyerabend 1981: 131-167.

—, 1970c/1999, "Experts in a Free Society", as reprinted in Feyerabend 1999: 112-126.

Feyerabend, Paul Karl, 1970d, "Against Method: Outline of an Anarchistic Theory of

Knowledge", in Radner, Michael, Stephen Winokur, eds., *Analyses of Theories and Methods of Physics and Psychology*, University of Minnesota Press, Minneapolis: 17-130.

—, 1975, *Against Method: Outline of an Anarchistic Theory of Knowledge*, New Left Books, London.

—, 1981, *Problems of Empiricism: Philosophical Papers, Volume 2*, Cambridge University Press, Cambridge.

—, 1999, *Knowledge, Science and Relativism: Philosophical Papers, Volume 3*, edited by John Preston, Cambridge University Press, Cambridge.

Finocchiaro, Maurice A., 2019, *On Trial for Reason: Science, Religion, and Culture in the Galileo Affair*, Oxford University Press, Oxford.

Galilei, Galileo, 1997, *Galileo on the World Systems: A New Abridged Translation and Guide*, edited by Maurice A. Finocchiaro, University of California Press, Berkeley.

Hesse, Mary B., 1967, "Galileo and the Conflict of Realism and Empiricism", in *Atti del Symposium Internazionale di Storia, Metodologia, Logica e Filosofia della Scienza*, Florence: 283-287.

Hopkins, Matthew, 1647, *The Discovery of Witches*, accessed online, May 2023, https://www.gutenberg.org/cache/epub/14015/pg14015.txt.

Kuhn, Thomas S., 1962, *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago.

Maier, Thomas, 2009, *Masters of Sex*, Basic Books, New York.

Masters, William H., Virginia E. Johnson, 1966, *Human Sexual Response*, Little, Brown and Company, Boston.

Merton, Robert K., 1969, "Behavior Patterns of Scientists", in *The American Scholar*, 38: 197-225.

Thayer, Henry S., 1953, ed., *Newton's Philosophy of Nature: Selections from his Writings*, Hafner Press, New York & London.

Wisdom, scientific expertise, and laypeople: some remarks on Paul Feyerabend's philosophy of expertise

Pierluigi Barrotta, Roberto Gronda

Abstract: Throughout his career, Feyerabend was seriously concerned with the authoritative role claimed by experts within democratic societies. He repeatedly argued that citizens should not be intimidated by the authority of science, and they should resist any attempt to strip themselves of their right to have a say in important social matters of public concern. We do not share Feyerabend's anarchist approach to philosophy of science; nevertheless, we believe that some aspects of his philosophy of science can easily be incorporated into a constructive philosophy of scientific expertise. The aim of this essay is to argue for two theses that we believe have an unequivocal Feyerabendian "flavour": a) that to be a good scientific expert, the scientist must be endowed with *wisdom*; and b) that public opinion is not limited to setting the goals that the scientific expert should take as exogenous data. In this way, we outline a normative model of the epistemic contributions that citizens and scientific experts can make to solve public problems.

Keywords: Feyerabend, Scientific expertise, Public problems, Separability thesis, Wisdom, Thick concepts, Citizens as epistemic contributors.

1. *Introduction*

Feyerabend had a very high opinion of the Founding Fathers of modern science, starting with Galileo, who managed to undermine established ways of thinking using all the means at his disposal, including rhetoric and inconclusive arguments. Feyerabend's admiration for modern science concerned its liberating effect from tradition. As he wrote: "in the 17th, 18th, even 19th centuries […] science was one of many competing ideologies […] In those years science was a liberating force, not because it had found the truth, or the right method […], but because it restricted the influence of other ideologies and thus gave the individual room for thought" (1978: 75).

The situation has radically changed today. Indeed, Feyerabend believed that "Ideologies can deteriorate and become dogmatic religion […]. They start deteriorating when they become successful, they turn into dogmas the moment the opposition is crushed: their triumph is their downfall. The development of

philinq XII, 1-2024, pp. 107-124
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.554
Submitted: 5 June 2024
Accepted: 22 July 2024

science in the 19th and 20th centuries, especially after the Second World War, is a good example" (1978: 75).

From these remarks we can understand the deep antipathy that Feyerabend had towards the presumed authority of scientific experts: "I think very highly of science, but I think very little of experts, although experts form about 95 percent or more of science today" (1970/1999: 112). Against the arrogance of scientific experts, Feyerabend suggests two remedies. The first concerns precisely scientific experts. In a free society, the scientific expert must go beyond the narrow scope of their specialization, open up to humanistic culture and become a *dilettante* (1970/1999: 117-118). The second concerns citizens. In a free society, citizens need not be intimidated by the authority of science: "science", Feyerabend writes, "is not beyond the reach of the natural shrewdness of the human race. I suggest that this shrewdness be useful to all important social matters which are now in the hand of experts" (1978: 98).

Feyerabend's suggestions appear naïve to us. Understanding the highly complex language of much of contemporary science seems to require a lot more than the "natural shrewdness of the human race". On the other hand, the idea that the scientific expert must be a *dilettante* is fascinating, but a little too vague. Even though we shouldn't give too much weight to the semantic origin of words, it is perhaps worth remembering that in Italian the word *dilettante* also has a pejorative meaning: it means a novice, a beginner or even an incompetent.

Nonetheless, we believe that the spirit of Feyerabend's suggestions should be accepted. Here we intend to defend two theses, which have an unequivocal Feyerabendian "flavor".

The first thesis is as follows: In order to be a good scientist, it is not enough to just be a good scientific expert. To be a good scientific expert, the scientist must be endowed with *wisdom*. The attribute of wisdom is not usually included among the qualities that a scientific expert must have. The first goal of our essay is to clarify the meaning of "wisdom" and show why it is necessary for the scientific expert.

The second thesis is as follows: public opinion is not limited to setting the goals that the scientific expert should take as exogenous data. Rather, in a well-functioning democracy, laypeople can and, in some cases, must a) help the scientific expert to better define socially relevant problems, and b) help the scientific expert, even epistemically, to solve socially relevant problems.

We will proceed this way. As for the first thesis, we will first briefly show, in Section 2, why it is not enough to have adequate scientific competence to be a good scientific expert: in other words, we will argue that we should distinguish between scientists and scientific experts. Second, in Section 3 we will clarify the concept of wisdom in the context of the philosophy of expertise. Then,

through the literature on wisdom, which is very extensive in moral philosophy, we will show how wisdom can perform different functions. To this literature, we will add a directly relevant topic for our discussion: each function of wisdom corresponds to a different figure of the scientific expert. In Section 4 we will tackle the second thesis. We begin by introducing the notion of "concerned citizen" and then go on to outline a set of desiderata that an account of citizens' contribution to inquiry must meet in order to be considered as genuinely epistemic. Finally, in Section 5 we will use the notion of thick scientific concepts to articulate the idea that citizens can act as semantic contributors by helping scientific experts and policymakers define the meanings of at least some of the concepts used in public inquiry.

2. *Scientists and scientific experts: the separability thesis*

We believe that there are good reasons to distinguish the activity of scientists, engaged in pure or basic research, from that of scientific experts, who instead have the task of helping the policymaker to solve problems of public importance. To understand these reasons, a brief historical review of the antecedents to the current debate on what is called the "separability thesis" between scientists and scientific experts is probably appropriate. It concerns a theme of central importance for evaluating Feyerabend's position.

According to the most traditional view, scientists can and should ignore moral values even when acting as consultants to a policymaker. The acceptability of the end, it is argued, is decided by the client using the scientist's experience. The latter, as a scientist, only has the task of evaluating the feasibility of the end and the means to obtain it. This is the ideology that Feyerabend complains about when he talks about the narrowness of scientists, incapable of opening up to humanistic culture.

After a seminal essay by Rudner (1953), the traditional conception was subjected to severe criticism within the sphere of analytical philosophy itself. The criticism is based on the "inductive risk" present in the *acceptance* of a theory: this not only depends on the empirical evidence but also on the importance, in a typically moral sense, that we give to the consequences of a wrong decision. For example, Rudner notes that marketing a medicine requires much more empirical support than marketing a machine for making buckles. The difference lies in the different importance we attribute to the possibility of erroneously accepting the two theoretical hypotheses underlying the decision to market the respective products. Rudner adds a second argument in favor of the thesis of the moral value-ladenness of scientific research, which was much less discussed by the several interventions that followed Rudner's essay (cf. Jeffrey

1956; Hempel 1965, 1981; Levi 1960; Lacey 2005; McMullin 1983).[1] Developing Quine's holistic theses against Carnap, Rudner states that in scientific research there is no way to separate practical choices from theoretical ones. As Rudner wrote, "I think that the statement that Scientists *qua* Scientists make value judgments, is also a consequence of Quine's position" (1953: 6).

Feyerabend certainly would not have accepted Rudner's proposal (not surprisingly, we are not aware that he ever made use of it). At the end of his essay, Rudner ventures to imagine a *science of ethics* to safeguard the objectivity of theoretical choices. A thesis that Feyerabend would certainly have branded as yet another case of chauvinism and imperialism of the scientistic mentality.

We have therefore reached the present day, where there has been a revival of Rudner's argument on inductive risk initially proposed by Douglas (2000; 2009). Douglas does not refer to holism to defend the thesis that all science is loaded with moral values. Rather, she prefers to refer to the undoubted authority of science in the contemporary world. Her argument is sociological, rather than epistemological. As Douglas writes: "Scientists hold a broadly authoritative position in our society, regardless of whether they are functioning in a formal advising role or not. Thus, when scientists make empirical claims, whether in scientific conferences, in science journals, or on an advisory panel, those empirical claims carry with them a prima facie authority [...]. The authority of science in society makes a distinction between scientist qua scientist and scientist qua advisor untenable" (2009: 82).

It is in all likelihood the sociological turn imposed by Douglas that suggested the thesis of the separability between the role played by scientists from that played by scientific experts. Mitchell (2004) and more markedly Gundersen (2018) have proposed a series of Merton-like behavioral norms that distinguish the activity of the scientist from that of the scientific expert engaged in the role of consultant.

We accept the separability thesis, even though we prefer to follow a different path. We will not delve into the topic here and will not attempt to justify the separability thesis from our perspective. It is sufficient to note that Gundersen does not appear convincing when he states that: "experts need not perform research to contribute to policymakers. Rather, they tend to apply already *existing knowledge* for the policy area in question" (2018: 56, italics added). There is no doubt that experts must apply the existing *theoretical* knowledge of one or more disciplines. However, experts must also acquire *new* knowledge to solve particular public problems and, most of the time, even unique in their complexity and characteristics. Theoretical knowledge must be articulated to

[1] Douglas (2009) provides an extensive overview.

take into consideration the social values at stake (sometimes, these values are expressed too vaguely to constitute the aim of a policy) and the physical particularities of the situation (knowledge of which is sometimes possessed only by the laypeople who live in the area; see, for instance, Wynne 1996; Barrotta *et al*. 2018). The role of scientific experts is, therefore, much richer and more *epistemically* sophisticated than how Gundersen depicts it. The scientist must demonstrate both epistemic and moral skills when she acts as a scientific consultant and expert.

In Section 3, we will develop this conclusion and introduce the concept of *wisdom* within the philosophy of expertise.

3. *Wisdom and scientific expertise*

The reference to the wisdom that scientists should have as experts may perhaps appear surprising (wisdom is not usually associated with scientific activities) or (symmetrically) an empty rhetorical device. In this paragraph, we intend to dispel these impressions.

The literature on wisdom has a long tradition connected to the Aristotelian concept of *phronesis*. In *Nicomachean Ethics*, it was Aristotle who emphasized that wisdom consists in knowing how to apply the "universal" to the "particular", to use his terminology (Aristotle 2002, especially Book VI). The Aristotelian conception of wisdom is therefore not only moral but also epistemic since it includes a good deal of cognitive ability to know whether the particular falls under the universal. The application and articulation of general theoretical knowledge are not a matter of deductive or mechanical activity. They require wisdom.[2]

If this is the case then being a good scientist, namely having reliable and updated general theoretical knowledge (the knowledge of laws, models, empirical uniformities, and the like), is a necessary but not a sufficient condition for being a good scientific expert. Scientific experts must also have the skill to apply and articulate general theoretical knowledge when facing particular situations.

This is the Aristotelian idea that we want to preserve in our discussion of the relationship between wisdom and scientific expertise. However, we will depart from the Aristotelian idea of wisdom from other points of view, which are not

[2] Although not the topic here, today various arguments support the thesis, very close to Aristotle's standpoint, that the explanation or prediction of a particular event is never strictly deducible from theoretical knowledge (along with initial conditions). Some of them could probably be traced back to John Stuart Mill (1843, especially Part III, Ch. XI), but they have been especially well clarified in one of Hempel's last works (1988). More recently, the point has also been raised by Nancy Cartwright (1999, especially where she criticizes the "received view"), albeit following a different line of reasoning, based on a different conception of scientific laws compared with that upheld by Hempel.

directly relevant here. For example, we will not deal with Aristotle's theory of emotions. Furthermore, another very important aspect of wisdom in Aristotle, which is not relevant for our purposes, is the connection between wisdom and happiness or eudaimonia. On the positive side, we are also interested in aspects that are absent in Aristotle. First of all, we would like to connect wisdom to public inquiries aimed at solving public problems. This is why in this paper we do not limit ourselves to the Aristotelian discussions of the concept of wisdom.

In the extensive literature on wisdom, we find that wisdom performs several functions. We will show how each function corresponds to a different figure of scientific experts: an interesting and so far unnoticed connection between moral philosophy and the philosophy of expertise.

We do not intend to offer an exhaustive list of the different functions that wisdom performs in public inquiries. The central point, important here, is that all these functions concern how scientific experts might apply their scientific competence to particular circumstances. These are the most important functions, which may overlap when addressing a particular public problem.

– *The instrumental function*. In this case, the purpose of the action is clearly defined. The expert only has to choose the most suitable means to achieve the goal. Although, in its concrete aspects, this function of wisdom very often requires delicate and complex evaluations, in its conceptual aspects it represents the simplest form of wisdom, since the goal set for the expert is already very clear and defined. In this case, the role of the expert is similar to that of the engineer. At least in its stylized form that we often find in literature, the engineer is the typical scientific expert who simply has to find in the given circumstances the most suitable means to achieve a well-defined end. She is not concerned with the end and the constitutive values underlying that end. Not surprisingly, it is the role of the expert that we find mainly in the empiricist or Humean tradition, where there is a clear-cut separation between ends or values, on the one hand, and technical-scientific evaluations, on the other. However, this is a very impoverished version of the function of wisdom and of the role that scientific experts often play. It is not by chance that in this context Aristotle does not use the word wisdom or *phronesis*, but rather the word cleverness or *deinotēs* (Russell 2009: 7, 24; Roberts *et al*. 2007: 306).
– *The function of specifying values*. Values are sometimes expressed too broadly to serve as an end (Russell 2009: 8-10; Richardson 1994: Ch. VIII), and this is especially true for the end of a policy. Take the welfare of a population as an example. To transform this value into an adequate goal, it is necessary to specify it better by carefully analyzing the preferences of the community. Or, take the environmental defense required by a community. Again, the

value can be specified in different ways to become the end of a policy. For example, the end could be implemented by preventing the urbanization of an area or by safeguarding green areas and minimizing pollution. Unlike the instrumental function, in this case the end is not given, since it must be carefully reconstructed by the expert starting from the particular situation in which she acts. It should be noted that laypeople's preferences are often given in a very confused way. In many circumstances, political decision-makers do not help at all to clarify the values at stake. Consequently, the task of scientific experts requires considerable skills in helping a community to transform values into specific ends.

– *The function of justifying the end*. This function of wisdom or practical intelligence has an obvious importance in ethics (Russell 2009: 7-11, 21-22; Richardson 1994) but it also plays an important role in scientific expertise. Perhaps for this reason, as a tribute to the empiricist and Humean tradition, the task that this function of wisdom plays in one of the possible roles performed by scientific experts has often been overlooked. However, this role is difficult to underestimate. For example, the goal of reducing carbon dioxide present in the atmosphere today is politically justified by climate experts who bring to light the negative consequences that a high percentage of carbon dioxide has on the climate. In cases like this, the scientific expert acts as an advocate for specific ends. Note that in these circumstances she violates the fact/value dichotomy and the alleged moral neutrality of scientific experts. However, she is not necessarily a bad expert. A scientific expert may act wisely by defending a certain purpose. More so than a layperson, it is the scientific expert who can understand the consequences of not achieving a certain end.
– *The function of handling conflicts among several ends*. One of the tasks assigned to wisdom in ethics literature concerns the best way to behave when two or more ends are incompatible in the given circumstances (Russell 2009: 25-26 and 31; Zagzebeski 1996: 221-224). Again, this function of wisdom provides a possible role for scientific experts. In many circumstances, public opinion and policymakers face real dilemmas. Conflicts of this kind are usually thought to be addressed through compromises, especially if different components of society support each end. However, scientific experts can have the role of showing that compromises are not necessary because there are technical and scientific tools that make compatible ends in apparent conflict. For example, at the beginning of the recent pandemic, the need to trace the movements of Covid patients was advocated. Many objected that traceability was incompatible with the value of privacy. However, scientific experts designed tracing methods that safeguarded both the need for health

and the value of privacy. Note that the role of the scientific expert as a reconciler between conflicting ends is also important when complete reconciliation is not at hand. Through this role, the scientific expert may be able to reduce the need for compromise, showing the partial compatibility of the ends. This happens regularly when, for example, the need for industrial development in an area clashes with the need to protect the environment.

– *The function of evaluating the overall means-end relationship*. Even if the means are adequate to achieve a given end, wisdom has the task of assessing whether the end achieved has unintended negative consequences. It is therefore the overall means-end relationship that must be evaluated. This function of wisdom or practical intelligence is underlined by Dewey's philosophy of values (cf. Dewey 1939), and clearly defines a further task of the scientific expert. To take a recent example, public opinion and policymakers may want a lowering of interest rates to favor businesses and families in difficulty. The aim is commendable, but the scientific expert might point out that, under certain circumstances, lowering interest rates leads to an increase in prices and therefore inflation. An end which, although positive, has further negative consequences is not necessarily a good end. This is why Dewey argues that often a good decision-maker should not limit herself to evaluating the means towards an end, but the overall means-end relationship.

Wisdom certainly has other functions and properties, some of which are not relevant to our discussion. For example, the wise person must not only know the reasons for a certain action but must also have the right psychological motivations to act consistently with those reasons. In any case, we believe it would be pointless to provide an exhaustive list (as long as an exhaustive list is possible). More interesting is to remark that every possible function of wisdom involves the ability to apply and articulate general theoretical knowledge to circumstances that are often unique and unrepeatable in all their details.

We hope we have managed to show the Feyerabendian "flavor" of this proposal. To be a good expert, the scientist should not take refuge in specialized knowledge. He should be able to understand the complexities of particular situations, including the social and moral values at stake. As we have shown, this requires a good deal of wisdom, a quality which today is rightly associated with humanistic culture and sensibility.

4. *The concerned citizen: framing the question*

To recall the general thesis from which we started, we argue that, in a well-functioning democracy, lay people can, and in some cases must, a) help the scientific expert to define socially relevant problems better, b) help the scien-

tific expert, even epistemically, to solve socially relevant problems. Now, the formation of the scientific expert – in the technical sense that we have just clarified – goes hand in hand with the formation of a particular kind of citizen who can actively participate in the process of public inquiry. We call this new kind of citizen "the concerned citizen".[3]

To remain faithful to the approach we have taken so far, we frame the question as follows: we ask what kind of epistemic function the concerned citizen is expected to perform and what kind of epistemic contribution they are expected to make in the context of public inquiry. It is important to note that we are concerned here exclusively with the *epistemic* nature of citizens' contribution. There is no doubt that citizens can make other relevant, non-epistemic contributions to public inquiry: for example, to mention only the most obvious one, they can provide political legitimacy to the decision reached through public deliberation. This is certainly an important aspect of the whole picture – one that should certainly not be underestimated if the aim is to develop a full and comprehensive account of citizens' public activity. But the recognition of the political role of citizens would hardly be a significant result, to say the least.

Our aim is, therefore, at the same time more limited and more ambitious: we want to show that citizens can act as epistemic agents when they are involved in a certain kind of public inquiry. We want to argue for something even stronger: we want to prove the stronger thesis that citizens can make an authentic epistemic contribution to public inquiry. For a contribution to be authentic, we assume that it must be proper to a certain group: in this sense, the contribution of scientific experts is authentic because they only have the skills necessary to generate relevant theoretical knowledge and apply it to the specific situation under consideration. Similarly, if citizens are to be viewed as genuine epistemic agents, they must be capable of making some epistemic contributions that scientific experts cannot make.

One possible way to account for such a contribution – which we are not going to pursue in this essay – would be to focus on the phenomenon of lay expertise or indigenous knowledge. Lay expertise refers to the fact that people who do not have any formal training or qualification – for instance, they do not have a degree – can nonetheless provide relevant pieces of information for solving the problem because of their rather unique experience of the situation: for example, the Cambrian farmers studied by Wynne (1996) were experts on both the composition of the soil and the behavior of their sheep because of their acquaintance with them.

[3] In this sense – and it is worth noting – our approach is openly and self-consciously normative, since we are trying to outline new social figures that could improve the overall quality of public problem-solving.

Now, the reason we don't want to go down that road is that lay expertise, though lay, is still a full-fledged form of expertise. It is a form of expertise that is acquired by being immersed in the situation rather than by undertaking a formal course of study. It also differs from scientific expertise in some other relevant respects: unlike scientific knowledge, for instance, lay knowledge does not claim to be universal or applicable to any context whatsoever. Its limits of validity are therefore quite narrow: they coincide with the limits of the experience made by lay experts. And yet, no one can reasonably deny that lay experts have relevant knowledge and that they should be involved in public inquiry.

Citizens or lay people are not lay experts. Or rather, it is not clear what kind of experience lay citizens could rely on to provide the equivalent of the local knowledge provided by lay experts. So, what other option is there available?[4] The idea is to shift the attention to the concepts through which a public problem is defined.

5. *Thick scientific concepts and public inquiries*

We begin by proposing a taxonomy. Concepts fall into many different categories, but for the purposes of our investigation, we focus on three that are particularly relevant.[5] Some concepts can be said to be purely scientific or tech-

[4] One might argue that citizens should be treated as experts in values, and that their lay expertise in values is the reason why they should be involved in public inquiry. Now, a reply along these lines is problematic for a number of reasons. First, it looks a bit like a trick. It is true that it is far from clear whether there are moral experts, let alone what such an expertise might look like. Nevertheless, it seems fair to say that if we want to use the term "experts" in its proper meaning, we should acknowledge that there must be some line separating experts from non-experts. If it is assumed that every citizen *qua* citizen is a moral expert, that amounts to a refutation of the very idea of moral expertise. One could try to reply that citizens' moral judgement is to be considered more authoritative than that of the experts and the professionals. But authoritative in which sense? If by "authoritative" we simply mean that in our societies decisions on values ultimately belong to the citizens, this is certainly true, but it does not show that the reasons why this is so are epistemic, namely, that citizens have a better and more reliable knowledge about values than experts and professionals. We still need an argument to prove this thesis. Anyway, we are not interested in taking a position on this point, as it is orthogonal to the issue we are discussing: if there is no moral expertise, citizens are not experts and, consequently, it is not clear how they should be able to give any relevant contribution to the inquiry. Even in the case in which citizens were considered moral experts, it would still be necessary to give an argument in support of the thesis that such an expertise could enable them to make an *epistemic* contribution to the inquiry.

[5] For instance, we do not consider here what may be called "commonsense concepts", namely, the set of concepts – factual or axiological – by which we conduct our daily business. Commonsense concepts are of the utmost importance in that they provide the source of semantic intuitions that citizens and laypeople use in every inquiry in which they participate. The reason why we decided to leave them aside is that they do not fit into the taxonomy, which is built around the thick/thin distinction. We do not consider it a problem since we do not see our taxonomy as anything close to a natural classification.

nical: "lepton" is a scientific concept in the sense that its meaning is fixed by the restricted community of physicists working on elementary particles. No one – we believe – could seriously argue that someone who is not a specialist in said field has the right to have a say on the meaning of "lepton". We are not interested in putting forth an account of why this is and should be the case: we pass the buck to those who reject this intuition.

Other concepts are thin, which also means that no group of people is recognized as having any exclusive and undisputed authority over their meanings. Examples of thin concepts are good, bad, fair, and many of the general ethical and axiological notions that we use to evaluate actions and performances. There might be arguments proving that there are or there should be moral experts – to name only the most structured debate – but our institutional settings as well as our conceptual practices do not align with that insight.

Still other concepts – which are the most interesting ones from our point of view – encompass scientific and axiological elements. We call them thick scientific concepts: thick scientific concepts are those for which the adoption of a normative standpoint is needed to fix the meaning of the concepts (Barrotta *et al.* 2024; on thick concepts, see Kirchin 2013). Some rather uncontroversial examples of thick concepts are those of happiness, well-being, disease, inequality, violence, or aggressiveness (see, for instance, Alexandrova 2012; Alexandrova *et al.* 2022; Keil *et al.* 2017; on the importance of thick concepts for social research, see Abend 2019). It has been argued that biodiversity can also be conceived of as a thick concept (Barrotta *et al.* 2020; see also Makineni *et al.* 2021). Risk and sustainability are other remarkable examples of thick scientific concepts (Möller 2012; Norton 2005).

Thick scientific concepts pose some serious challenges to the received view of value-free science. It is not just that scientists qua scientist make value judgments to accept or reject a certain hypothesis; more radically, the very concepts that they use in their inquiries are value-laden through and through. While there may be some serious philosophical debate about their nature, however, there is no doubt that thick scientific concepts can be made objects of rigorous scientific investigations. The economy of happiness or the economy of well-being, not to mention the numerous different approaches to the study of human behavior or the theories of risk assessment, are well-established fields of research.

At the same time, these concepts play an important role in public policy. And while their content is – at least in part – determined by scientific inquiry, there is a sense in which their meaning also depends on how citizens and lay people use these concepts to understand their actions and social interactions. Take the concept of happiness, for example: everyone – we guess – has an opinion

about what a happy life should be like, and it would be hard to argue that their semantic understanding is completely irrelevant to the question of determining what happiness means – especially when it comes to an understanding of what public happiness or the happiness of the public should look like.[6]

It is important to stress the difference between scientific and public inquiries. To the best of our knowledge, a comprehensive and satisfactory theory of inquiries is still lacking, although the topic has recently received more and more attention, especially in the epistemological debate (Friedman 2023; Kelp 2021). For our purposes, however, some general and preliminary remarks on the structure of inquiries will be enough. We assume that scientific inquiries are those whose criteria of success are set and judged by the community of scientists; public inquiries, on the other hand, are those whose criteria of success are to be set by the larger community, which includes scientific experts, citizens, politicians, stakeholders, etc. This means that it is up to the wider community of inquirers to decide how the problem should be defined. Accordingly, while it can be argued that in scientific inquiry scientists should be free to define thick scientific concepts as they wish – for example, by choosing whatever indicators they feel are appropriate to the task – public inquiry, which aims to solve real-world problems, requires a different attitude towards meaning determination.

So, for instance, the public inquiry of ensuring the well-being of a community cannot come out with a solution that meets the standards of a certain discipline but is completely unsatisfactory to the members of the public. The

[6] Two remarks can be useful here. First, we don't mean to suggest that a person or a group of people cannot be wrong about the meaning that they ascribe to a word of a concept: if someone said that the greatest happiness in life is being beaten to death, we would assume that they were either mad or didn't properly grasp the meaning of happiness. We are not advocating semantic infallibilism. Our point is rather that a very strong argument is needed to support the idea that the meaning of thick scientific concepts such as happiness or well-being – when they are used in the context of policy-making – should be fixed and determined by the experts alone, and that citizens and laypeople do not have right to have a say on them. And we don't see how such an argument can be made unless you are prepared to buy into the idea of scientism in its strongest version possible, namely, the idea that "science" is the measure of all things, even of the meanings that laypeople use to understand their own place in the world. Secondly, we don't want to limit the scientists' freedom of research and their right to self-determination: of course, a team of scientists can freely decide which proxies or indexes to choose to measure happiness, well-being, and so on, and operationalization is one of the ways in which the meaning of a concept is fixed, or explicated if you prefer. Nonetheless, there should intuitively be some constraints on the choice they are allowed make if their aim is to understand the phenomenon (or some aspects of the phenomenon) that goes under that name. If, for instance, the only indicator considered when measuring the well-being of a population were their average height, we would be hesitant to say that such a measurement has to do with well-being. We would probably say that they are using that concept in a somehow idiosyncratic way or that we are talking past each other because we are referring to two different things.

same point can be made in a slightly different way, by saying that the object to which the public problem refers cannot be reduced to one of its possible components: our favorite example is the decision to build a dam in a certain place. The intuition that we would like to capture is that such a problem is not merely an engineering problem or an economic problem or a hydrogeological problem or a real estate problem. When properly conceived a public problem is all these problems and many others taken together (Barrotta *et al.* 2022).

From these observations and remarks, we derive our semantic conclusions. Our thesis is that when thick scientific concepts are used in a public inquiry, their meaning is open to negotiation: citizens, as well as scientific experts, stakeholders, and policymakers, have the right to have a say on what those concepts mean in that particular context. One risk that must be avoided is that the meaning of thick scientific concepts be fixed in purely disciplinary or technical terms, thus betraying the publicity of the public inquiry. We guess we are here as close as we can get, moving from our background, which is strongly indebted to classical pragmatism, to Feyerabend's agenda of removing the obstacles that intellectuals and specialists erect to prevent laypeople from questioning their authoritative role.

One genuine *semantic* contribution that citizens can make to the public inquiry – we don't want to rule out, in principle, the possibility that other contributions can be singled out – is, therefore, to help clarify and define the thick scientific concepts that are used in the course of inquiry. Such a contribution is genuine in the sense that it cannot be made by any of the other groups participating in the public inquiry: citizens – who are in most cases not only the immediate or ultimate object of public inquiry, but also one of the main drivers of change when it comes to public inquiries – may have, and usually have, a specific understanding of what those concepts mean which is grounded on their experiences, social standpoints, private or community interests. In other words, the contribution that concerned citizens can make to public inquiry is to ensure that the specificity of the situation is properly taken into account by the scientific experts who are called in to solve the public problem by resisting any use of the (thick scientific) concepts that goes against their understanding of them. In Habermasian jargon, this means that citizens can help scientific experts exercise their wisdom by rejecting any forms of colonization of their life-worlds that scientific experts may more or less inadvertently introduce.

To conclude our argument, we would like to add a few clarifications that can be useful to avoid potential confusion. Firstly, at this stage we are not ready yet to put forth a stronger thesis, to the effect that thick scientific concepts are either a necessary or a sufficient component of a public inquiry *qua* public. We strive to get there – we believe that their presence in an inquiry could be stated

as a necessary condition for its publicity – but we don't have enough material to back up that claim: we therefore rest satisfied with a moderate and pluralistic approach, according to which that of making a semantic contribution is one of the functions that citizens can perform within public inquiry.

Secondly, it may be worth adding a few more words about the kind of constraints that the understanding that citizens have about the meaning of thick scientific concepts imposes on public inquiry. We do not want to argue that the meaning that citizens attribute to a thick scientific concept is to be accepted as a fact by all the other participants in the inquiry. That would entail a form of infallibilism that is at odds with the self-corrective nature of inquiry. Our point is rather to suggest that by participating as genuine contributors to the inquiry citizens may undertake a process of semantic self-clarification at the end of which they can eventually revise or even reject their original understanding of the concepts used within the inquiry. Such a process of semantic revision is made possible by, and goes hand in hand with, the wise application of scientific knowledge to the specific conditions under consideration made by scientific experts. Our ultimate goal, therefore, is not to make an argument for replacing scientists with citizens at the center of society – to use, once again, Feyerabend's formula – but to argue for the (epistemic) necessity of enlarging and enriching the community of inquirers by including citizens and non-experts.

6. *Conclusion*

In this essay we have tried to show that despite his self-professed anarchism and dadaism, some aspects of Paul Feyerabend's philosophy of science can easily be incorporated into a constructive philosophy of scientific expertise. In particular, we have relied on the theoretical resources made available by traditions of thought as different as American pragmatism (broadly conceived) and an Aristotelian-inspired account of wisdom to translate Feyerabend's reasons of distrust in scientific experts into a set of requirements that citizens and scientific experts must meet in order to participate in the public life of a well-functioning democracy.

Our analysis has resulted in a normative model of the different epistemic functions that laypeople and scientific experts can be expected to perform to contribute to solving a public problem. We take it as a way of preserving an important Feyerabendian thesis, namely, that science – or, rather, scientists, since science is nothing more and nothing less than what scientists do as a group – should not be allowed to decide alone on matters of public importance. In this sense, Feyerabend's plea for a proliferation of points of view is still a remarkable and valuable insight: adding voices to the public debate is certainly useful

to keep at bay the more or less conscious temptation of scientists to gain total control over the public sphere.

However, we have also argued that something more needs to be added to this rather minimal insight. By focusing on the epistemic contributions that the scientific expert and the concerned citizen can make to public inquiries – from the very definition of the public problems to the evaluation of the different possible plans of action – we have tried to show that there are sound epistemic reasons for enlarging the community of inquirers, thus outlining an epistemic justification for citizen participation in the activities of public problem-solving.

There is a point, however, about which Feyerabend is unquestionably right. In our democratic societies, the use of scientific knowledge to solve socially relevant problems is far from being flawless: public inquiries often tend to slide either towards technocratic solutions or towards a general and chaotic involvement of citizens, with no interest in the epistemic outcomes of their participation. Both alternatives are deeply unsatisfactory. In this sense, our model is self-consciously normative: we are aware that scientific experts (as we understand them) and concerned citizens are not well-established roles in our democratic societies. The importance of these figures for the public sphere can be defended and advocated from a philosophical point of view; however, their actual existence depends only on the decision of the public to educate and train its members to perform those functions that are needed to successfully carry out inquiries aimed at solving socially relevant problems.

Pierluigi Barrotta
University of Pisa
pierluigi.barrotta@unipi.it

Roberto Gronda
University of Pisa
roberto.gronda@unipi.it

Author's contribution: The Authors contributed equally to this paper.

Acknowledgments: This work was financially supported by the ISEED (Inclusive Science and European Democracies) project, funded from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 960366.

References

Abend, Gabriel, 2019, "Thick Concepts and Sociological Research", in *Sociological Theory*, 37, 3: 209-233.

Alexandrova, Anna, 2012, "Well-Being as an Object of Science", in *Philosophy of Science*, 79, 5: 678-689.

—, Mark Fabian, 2022, "Democratising Measurement: or Why Thick Concepts Call for Coproduction", in *European Journal for Philosophy of Science*, 12, 7: 1-23.

Aristotle, 2002, *Nicomachean Ethics*, translated and edited by Roger Crisp, Oxford University Press, Oxford.

Barrotta, Pierluigi, Roberto Gronda, 2020, "What is the Meaning of Biodiversity? A Pragmatist Approach to an Intrinsically Interdisciplinary Concept", in Allwood, James S., Olga Pombo, Clara Renna, Giovanni Scarafile, eds., *Controversies and Interdisciplinarity: Beyond Disciplinary Fragmentation for a New Knowledge Model*, John Benjamins, Amsterdam: 115-131.

Barrotta, Pierluigi, Roberto Gronda, 2022, "Intelligence and Scientific Expertise", in *Synthese*, 200: 142.

—, 2024, "Thick Scientific Concepts and Public Problems: Citizens as Semantic Contributors", in Janack, Marianne, Giancarlo Marchetti, Roberto Gronda, eds., *Pragmatism and Philosophy of Science: Historical and Contemporary Perspectives*, Routledge, London-New York: forthcoming.

Barrotta, Pierluigi, Eleonora Montuschi, 2018, "The Dam Project: Who Are the Experts? A Philosophical Lesson from the Vajont Disaster", in Barrotta, Pierluigi, Giovanni Scarafile, eds., *Science and Democracy: Controversies and Conflicts*, John Benjamins, Amsterdam: 18-33.

Cartwright, Nancy, 1999, *The Dappled World. A Study of the Boundaries of Science*, Cambridge University Press, Cambridge.

Dewey, John, 1939, *Theory of Valuation*, in Dewey, John, *The Collected Works*, edited by Jo Ann Boydstone, 1969-1991, *The Later Works*, Volume 13: 63-188.

Douglas, Heather, 2000, "Inductive Risks and Value in Science", in *Philosophy of Science*, 67: 559-579.

—, 2009, *Science, Policy, and the Value-free Ideal*, University of Pittsburgh Press, Pittsburgh.

Feyerabend, Paul Karl, 1970/1999, "Experts in a Free Society", as reprinted in Feyerabend, Paul K., 1999, *Knowledge, Science and Relativism: Philosophical Papers, Volume 3*, edited by John Preston, Cambridge University Press, Cambridge: 112-126.

—, 1978, *Science in a Free Society*, New Left Books, London.

Friedman, Jane, 2023, "The Aim of Inquiry?", in *Philosophy and Phenomenological Research*, Early View: 1-18.

Gundersen, Torbjørn, 2018, "Scientists as experts: A distinct role?", in *Studies in History and Philosophy of Science Part A*, 69: 52-59.

Hempel, Carl Gustav, 1965, "Science and Human Values", in Hempel, Carl Gustav, *Aspects of Scientific Explanation*, The Free Press, New York: 81-96.

—, 1981, "Turns in the Evolution of the Problem of Induction", in *Synthese*, 46: 389-404.

—, 1988, "Provisoes: A Problem Concerning the Inferential Function of Scientific Theories", in *Erkenntnis*, 28, 2: 147-164.

Jeffrey, Richard, 1956, "Valuation and Acceptance of Scientific Hypotheses", in *Philosophy of Science*, 22: 197-217.

Keil, Geert, Ralf Stoecker, 2017, "Disease as a Vague and Thick Cluster Concept", in Keil, Geert, Lara Keuck, Rico Hauswald, eds., *Vagueness in Phychiatry*, Oxford University Press, Oxford: 46-74.

Kelp, Christoph, 2021, "Theory of Inquiry", in *Philosophy and Phenomenological Research*, 103, 2: 359-384.

Kirchin, Simon, 2013, ed., *Thick Concepts*, Oxford University Press, Oxford.

Lacey, Hugh, 2005, *Is Science Value Free? Values and Scientific Understanding*, 1st ed. 1999, Routledge, London and New York.

Levi, Isaac, 1960, "Must the Scientists Make Value Judgments?", in *The Journal of Philosophy*, 57: 345-357.

Makineni, Vamsi, Sahotra Sarkar, 2021, "'Biodiversity' as a Primarily Normative and Inseparable Thick Concept", http://philsci-archive.pitt.edu/19562/.

McMullin, Ernan, 1983, "Values in Science", in Asquith, Peter D., Thomas Nickles, eds., *Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association*, Philosophy of Science Association, East Lansing, vol. 1: 3-28.

Mill, John Stuart, 1843, *A System of Logic*, 1967 reprint, Longmans, London.

Mitchell, Sandra, 2004, "The Prescribed and Proscribed Value in Science Policy", in Machamer, Peter, Gereon Wolters, eds., *Science, Values, and Objectivity*, University of Pittsburgh Press, Pittsburgh: 245-255.

Möller, Niklas, 2012, "The Concepts of Risk and Safety", in Roeser, Sabine, Rafaela Hillerbrand, Per Sandin, Martin Peterson, eds., *Handbook of Risk Theory. Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, Springer, Cham: 55-85.

Norton, Bryan, 2005, *Sustainability: A Philosophy of Adaptive Ecosystem Management*, The University of Chicago Press, Chicago.

Richardson, Henry S., 1994, *Practical Reasoning about Final Ends*, Cambridge University Press, Cambridge.

Roberts, Robert C., W. Jay Wood, 2007, *Intellectual Virtues. An Essay in Regulative Epistemology*, Clarendon Press, Oxford.

Rudner, Richard, 1953, "The Scientist *qua* Scientist Makes Value Judgments", in *Philosophy of Science*, 20: 1-6.

Russell, Daniel C., 2009, *Practical Intelligence and the Virtues*, Clarendon Press, Oxford.

Wynne, Brian, 1996, "May the Sheep Safely Graze? A Reflexive View of Experts-Lay Knowledge Divide", in Lash, Scott, Bronislaw Szerszynski, Brian Wynne, eds., *Risk, Environment and Modernity*, Sage, London: 44-83.

Zagzebski, Linda, 1996, *Virtues and the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundation of Knowledge*, Cambridge University Press, Cambridge.

Feyerabend's humanitarian pluralism and its relevance for science-based policy

Karim Bschir

Abstract: A strong commitment to pluralism on multiple levels (methodological, theoretical, ontological as well as political) is a defining feature of Paul Feyerabend's philosophical corpus. However, for Feyerabend, pluralism is not just an epistemologically preferable account within the philosophy of science. He also believes that pluralism is the only account of science that is compatible with a humanitarian outlook.

In the first part of this paper, I will reconstruct Feyerabend's theoretical pluralism in the context of his criticism of Thomas Kuhn's account. I will show that Feyerabend's critical engagement with Kuhn's model of scientific revolutions in the early 1960s was crucially important for the development of his own pluralistic account of science. In the second part, I will discuss and critically analyse the ethical-political stance that underlies Feyerabend's pluralism. In the final part, I briefly summarize a series of papers that I have published together with Simon Lohse, in which we apply Feyerabend's pluralism to current discussions about the role of evidence-based policy advice during the COVID-19 pandemic.

Keywords: Feyerabend, Criticism of monism, Theoretical and methodological pluralism, Humanitarianism, Science-based policy.

1. *Introduction*

It is impossible to pin down Paul Feyerabend's thinking to one central claim, position, or any sort of philosophical "ism". There are, however, certain recurring topics and motifs which play important roles throughout Feyerabend's intellectual trajectory.

One of these central topics is the idea of pluralism and its relevance for scientific knowledge production, which appears very early in Feyerabend's writings. His interest in philosophical debates on quantum mechanics in the 1950s, particularly his engagement with the positions of David Bohm and Niels Bohr, was highly influential for the arguments for pluralism that he would extend and further develop over the next decades.[1] Another important influence on

[1] For the influence of Bohm and Bohr on Feyerabend's early theoretical pluralism see van Strien (2020). See also Kuby (2021).

philinq XII, 1-2024, pp. 125-144
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.555
Submitted: 7 April 2024
Accepted: 22 July 2024

Feyerabend's early pluralism was Karl Popper, whom he first met in 1948 and whose lectures on critical rationalism he attended in the early 1950s at the London School of Economics. Popper's critical methodology as well as his emphasis on the function of theoretical alternatives in theory testing would become a key element of Feyerabend's own epistemic pluralism. Feyerabend typically framed his pluralist arguments in the context of criticisms of contemporary positions in the philosophy of science of his time. In this regard, his criticism of Thomas Kuhn's notion of normal science, as laid out in *The Structure of Scientific Revolutions* (1962), constituted a crucial step towards the development of his pluralism.

Another important recurring element in Feyerabend's philosophy is his merging of epistemological arguments with ethical-normative considerations. From a very early point in his career, Feyerabend combined philosophical criticisms, for example of logical empiricism or of Kuhn, with allegations of dogmatism, authoritarianism, and even tyranny. Especially in his later writings, and particularly in his 1978 book *Science in a Free Society* (which resembles the more famous book *Against Method* in that it is a collage of ideas and elements of texts that had been used in earlier publications), the ethical stance underpinning many of his epistemological arguments becomes more accentuated. For Feyerabend, pluralism is not just a position that is favorable compared to other accounts in the philosophy of science; he considers it to be the only account of science that is compatible with a humanitarian outlook and a free and democratic society in which all traditions and individuals can develop freely. Feyerabend's systematic arguments for pluralism as a preferable position within the philosophy of science simply cannot be separated from his moral and political defenses of pluralism as a basis for a free and open society.

However, the relationship between Feyerabend's pluralism and his humanitarianism (or his allegations of anti-humanitarianism towards the positions that he criticizes) is an underexplored aspect of Feyerabend scholarship. This paper aims to provide a first step towards closing this gap in the secondary literature.

I will begin, in Section 2, by describing the development of Feyerabend's early theoretical pluralism against the background of his criticism of Kuhn. The picture that will emerge from this will be incomplete. There are many more influences one would have to consider in a full historical reconstruction of the development of Feyerabend's pluralistic account. However, the specific points Feyerabend raises in his critical engagement with Kuhn's model nicely highlight those aspects in the general philosophical accounts of his time that he took issue with. His attacks on Kuhn's position also illustrate how he was eager to underline the function of criticism in science and the importance of fallibil-

ism as an essential epistemological attitude, which he had encountered and appreciated in Popper's philosophy and, later, in John Stuart Mill's *On Liberty*.[2]

In Section 3, I will turn to Feyerabend's humanistic defense of pluralism. As it turns out, Feyerabend never explicitly explains what he means by "humanitarianism" and why he thinks that accounts that allow for mild forms of dogmatism in scientific knowledge production (such as Kuhn's or others) are supposedly anti-humanitarian. I will suggest that Feyerabend's invocations of humanitarianism should be interpreted in the context of a more general critique of scientific modernity.

In Section 4, I will address whether Feyerabend's frequent invocations of humanitarianism can be understood as the expression of a genuine ethical stance. I will show that, if anything, Feyerabend implicitly proposes a particularist ethics in combination with a voluntarist approach to epistemology. Ultimately, it remains a matter of interpretation what Feyerabend's ethical stance exactly consists of, but it is possible to extract elements of such a stance from his writings.

The fact that Feyerabend does not articulate a systematically sound normative foundation for his pluralism does not mean that his pluralistic account of science is without merit. Quite on the contrary, the latter may be seen as relevant for current discussions about science and about the ways in which scientific knowledge can and should be implemented in society. To illustrate this, I will summarize, in Section 5, the main insights of previous collaborative work with Simon Lohse, in which we use a charitable reconstruction of Feyerabend's pluralism to deepen our understanding of specific shortcomings in the scientific policy advice during the COVID-19 pandemic.

2. *Feyerabend's early epistemic pluralism*

Feyerabend's early epistemic pluralism is best interpreted in a negative way: as a rejection of theoretical monism understood as the shared commitment of a scientific community to a single theoretical paradigm. Such a notion of monism plays a prominent role not only in the logical empiricist tradition of the early 20th century but also in Thomas Kuhn's theory of scientific revolutions and his concept of normal science.

Feyerabend's critical engagement with empiricism and Kuhn's theory in the early 1960s played a crucial role in the development of his pluralistic position. In 1961, Feyerabend received a copy of the manuscript of Kuhn's *The Structure of Scientific Revolutions*. Feyerabend wrote four extensive letters to Kuhn with detailed comments on the manuscript. The letters contain a multifaceted

[2] See Lloyd (1997) for the similarities in Feyerabend and Mill.

critique of Kuhn's account in *Structure*.[3] From 1961 onwards, the arguments used in his critique of Kuhn appear frequently in Feyerabend's writings including his contribution to the 1970 Lakatos/Musgrave volume (Feyerabend 1970), and in his magnum opus *Against Method* (1975b).

According to Kuhn's model of scientific change, phases of normal science are punctuated with scientific revolutions that lead to radical changes in the dominating paradigm of a discipline. During a normal scientific phase, the discipline is dominated by a single paradigm whose underlying theoretical assumptions are not questioned. Furthermore, normal science is characterized by a consensus about what counts as relevant and solvable problems within a discipline. Normal science thus has an important function in the development of science because adherence to a single paradigm enables scientists to select and focus on relevant problems, allowing for cumulative progress. A normal scientific phase ends when only those problems remain that cannot be solved by the problem-solving strategies of the dominating paradigm. Kuhn calls these problems anomalies. In Kuhn's view, the emergence of anomalies is the natural result of normal scientific activity.

It is at this point that Feyerabend's criticism of Kuhn sets in. Feyerabend believes that scientists should not just adhere to the prevailing paradigm, but actively seek to develop and entertain incompatible alternatives. According to Feyerabend, the availability of such alternatives can enhance the visibility of anomalies that would otherwise remain hidden.[4]

Feyerabend thus contrasts the monism that is characteristic of Kuhn's normal science with a notion of epistemic pluralism containing two main features. The first is the idea of the proliferation of theoretical alternatives. The second is the idea that descriptions of empirical facts, which potentially constitute anomalies, are dependent on the theoretical framework in which they are described and that the number of potential anomalies increases if multiple theoretical frameworks are used to describe empirical reality:

> [A]lternatives are both used and needed; and they are needed as it is only with their help that it is possible to find anomalies in whatever theory is being held at a special moment (Feyerabend to Kuhn in Hoyningen-Huene 1995: 366).

In a nutshell, Feyerabend's early epistemic pluralism consists in the methodological prescription to increase the likelihood of detecting anomalies in established theories. Alternatives are often necessary, according to Feyerabend,

[3] The letters are reproduced in Hoyningen-Huene (1995; 2006).

[4] For details on Feyerabend's criticism of Kuhn see Hoyningen-Huene (1995; 2000; 2006), Sirtes *et al.* (2006), Bschir (2015), Gillies (2024).

to detect and highlight the limitations and errors of a given theory or approach. The idea that alternatives increase the empirical content of established theories, thus enhancing their testability, also plays a central role in *Against Method*:

> Not only is the description of every single fact dependent on *some* theory [...], but there also exist facts which cannot be unearthed except with the help of alternatives to the theory to be tested, and which become unavailable as soon as such alternatives are excluded. This suggests that the methodological unit to which we must refer when discussing questions of test and empirical content is constituted by a *whole set of partly overlapping, factually adequate, but mutually inconsistent theories* (Feyerabend 1975b: 39).

If progress through paradigm change is the aim of science, then the proliferation of theories becomes a methodological imperative for science.

> Hence, if change of paradigms is our aim then we must be prepared to introduce and articulate alternatives [...], we must be prepared to accept a *principle of proliferation*. Proceeding in accordance with such a principle is one method of precipitating revolutions (Feyerabend 1970: 205).

In its early phase, Feyerabend's pluralism can be characterized as a systematically argued position that grew out of a critical engagement with the main accounts in the philosophy of science of the time. As such, Feyerabend introduced pluralism mainly for the sake of its epistemological benefits and defended it based on epistemological arguments. These arguments build on the theory-ladenness of experience and empirical descriptions, the importance of alternatives in theory testing, the importance of criticism in science, and a fallibilist stance according to which even the best scientific theories are likely to fail in solving certain empirical problems.[5]

It is important to note, however, that already from very early on, Feyerabend supported his pluralistic position not only with systematic epistemological arguments but also with ethical-normative reasons. These normative defenses of pluralism were intricately connected with the notion that pluralism is also a desirable political and social goal in a free society.

3. *Feyerabend's ethical-normative defense of pluralism*

As early as 1961, Feyerabend clearly expressed concerns about monism from both an epistemological and a moral perspective. In a quite obscure and

[5] For details on Feyerabend's theoretical pluralism see Martin (1972), Preston (1997), Lloyd (1997), Farrell (2003), Oberheim (2006), Sirtes *et al.* (2006), Bschir (2015), Tambolo (2015), Collodel (2016).

rather provocative passage, he associates Kuhn's historical analysis with an ideology that he clearly rejects. He compares Kuhn to those who point to history to justify their crimes. And he accuses Kuhn of being irrational:

> Your hidden predilection for monism (for one paradigm) leads you to a false report of historical events. You regard as one paradigm (classical physics, for example) which is in fact a bundle of alternatives […]. Which only confirms what I have said on the first page, viz. that you do not write history plain and simple, but that you present an ideology, and a very questionable monolithic ideology at that in the covers of history. In this respect you are really very similar to those who point to history in order to justify their crimes. You are a mystic, an irrationalist (Feyerabend to Kuhn in Hoyningen-Huene 1995: 367).

In his 1963 paper "How to be a Good Empiricist: A Plea for Tolerance in Matters Epistemological", we find one of the first associations of pluralism with what Feyerabend calls a "humanitarian outlook":

> *Unanimity of opinion may be fitting for a church, for the frightened victims of some* (ancient*, or modern) myth, or for the weak and willing followers of some tyrant; variety of opinion is a feature necessary for objective knowledge; and a method that encourages variety is also the only method that is compatible with a humanitarian outlook* (1963/1999: 97).

Feyerabend repeats this line of argument in the 1970 volume edited by Imre Lakatos and Alan Musgrave, which unites discussions of Kuhn's *Structure of Scientific Revolutions*:

> This ideology [which forms the background of Kuhn's thinking], so it seemed to me, could only give comfort to the most narrowminded and the most conceited kind of specialism. It would tend to inhibit the advancement of knowledge. And it is bound to increase the anti-humanitarian tendencies which are such a disquieting feature of much of post-Newtonian science (1970: 197-198).

And he reiterates the incompatibility of normal science with humanitarianism:

> This, I regard as the final and the most important argument against a "mature" science as described by Kuhn. Such an enterprise is not only ill-conceived and nonexistent; its defence is also incompatible with a humanitarian outlook (1970: 210).

That Feyerabend's epistemological arguments for pluralism also have implications beyond the philosophy of science and that a defense of pluralism for Feyerabend also has political ramifications, becomes clear in the 1965 essay "Problems of Empiricism", in which he associates "the idea of absolute knowledge" (which allegedly lies behind the monisms of Kuhn and the empiricist tradition) with tyranny and the "multiplicity of ideas" with democracy:

Is not a tyranny the natural correlate of the idea of absolute knowledge, and is not indoctrination the method of teaching most appropriate to it? Conversely, is not the idea of fallibility and the correlated demand for a multiplicity of ideas, and the hope that truth will arise from the civilized clash of such ideas the essence of all democracy? (1965: 217).

Feyerabend emphasizes the association of pluralism with a "humanistic system of belief" in another passage of the same essay:

All this shows that theoretical monism is the reflection, within the domain of theoretical knowledge, of a *much more general point of view*, which […] left its traces in almost all human activities. Conversely, we may guess that the theoretical pluralism that we propagate in epistemology may lend itself to generalization and may then lead to an outlook in the arts and in religion, as well as to a new, comprehensive ideology that assembles the scattered remains of a long-forgotten tribal ideology and unites them in a truly humanistic system of belief (1965: 218).

Feyerabend seems convinced that monism is not just an epistemologically deficient position within the theory of science but that it reflects a general viewpoint that permeates all aspects of society. He associates this viewpoint with dogmatism, indoctrination, tyranny, authoritarianism, hegemony, and narrow-mindedness; and he sees it as detrimental to the free development of individuals. While he clearly asserts that monistic philosophical accounts are incompatible with humanitarianism, he does not provide a systematic argument for this claim. But he makes it unmistakably clear that opposing views to pluralism are morally reprehensible and must be rejected on normative grounds.

Simultaneously, Feyerabend presents his own pluralism as morally virtuous. In the Introduction to the Chinese Edition of *Against Method*, he states that his main goal was humanitarian rather than intellectual:

My main motive in writing the book was humanitarian, not intellectual. I wanted to support people, not to "advance knowledge". […] Today old traditions are being revived and people try again to adapt their lives to the ideas of their ancestors. I have tried to show […] that science, properly understood, has no argument against such a procedure. There are many scientists who act accordingly. […] I am not against a science so understood. Such a science is one of the most wonderful inventions of the human mind. But I am against ideologies that use the name of science for cultural murder (Introduction to the Chinese edition, reprinted in Feyerabend 1993: 3-4).

In this passage, it becomes clear that Feyerabend's defense of pluralism is embedded in a specific and rather idiosyncratic cultural critique. He attacks Western science and the hegemonic ambitions that he associates with the pro-

ponents of the Western rationalistic-scientific worldview because he believes that it leads to a diminishment of the affluence and diversity of human cultures and non-scientific traditions.[6]

In a less considered contribution to a German volume in memory of the Austrian philosopher Walther Schmied-Kowarzik, who was Feyerabend's high school teacher during the war, Feyerabend identifies the more general phenomenon opposing pluralism with the expansion of technological-scientific civilization. He writes:

> One of the most important and most depressing phenomena of our time (since the 19th century) is the constant expansion of the Western technological-scientific civilization, its ways of thinking, its problem framings, and the destruction of cultures associated with its expansion. […] [T]he underlying phenomenon is an increasing monotony under the banner of an unholy alliance between business, technology, and primitive fragments of the sciences (1985: 138, my translation).[7]

This passage reflects his criticism of the tradition of Western science and rationalism that had already been put forward in a rather polemic fashion in *Science in a Free Society*. There, Feyerabend argued that science and the associated materialist-rationalist worldview constitute just one of many traditions, and in a free society, all traditions, including non-scientific and non-rationalist ones, should "*have equal rights and equal access to the centers of power*" (1978: 9).[8]

Feyerabend's extension of pluralism beyond the realm of science and epistemology is part of a general critique of scientific modernity. At the core of Feyerabend's critique stands the fear of uniformity and the hegemony of a scientistic technological worldview that would displace non-Western ways of life and diminish the abundance of life and human experience. Kidd (2021) identifies three features of Feyerabend's critique of scientific modernity: 1) an

[6] For Feyerabend's critique of scientific rationality see also: Feyerabend (1975a; 1976; 1978; 1980; 1987; 2011).

[7] The German original reads as follows: "Eines der wichtigsten und deprimierendsten Phänomene unserer Zeit (und schon des 19. Jahrhunderts) ist die ständige Ausdehnung der westlichen technologisch/wissenschaftlichen Zivilisation, ihrer Denkweise, ihrer Fragestellungen und die mit dieser Ausbreitung verbundene Zerstörung von Kulturen. […] [D]as grundlegende Phänomen ist eine zunehmende Monotonie unter dem Banner einer unheiligen Allianz von Business, Technologie und primitiver Bruchstücke der Wissenschaften". Similar statements can be found in many publications form the 1970s onwards.

[8] Feyerabend never provides a strict definition of the term "tradition". However, it becomes clear from his writing in *Science in a Free Society* that a tradition may be seen as a specific way of viewing/approaching life and reality. Thus, traditions clearly contain an element of what is often called "worldview" including the specific forms of knowledge that come also with that worldview. But traditions in Feyerabend's sense clearly also have a practical dimension. They are also ways of life. See also Farrell (2003) for a discussion Feyerabend's use of terms like "tradition" and others.

element of epistemic, environmental, and cultural violence; 2) a form of philistine scientism that goes along with the devaluation of the arts and humanistic values in comparison with the usefulness of science, engineering, and technology; 3) an existential disenchantment of human life due to an impoverished view of reality and nature provided by modern science.

While Kidd acknowledges the merits of Feyerabend's critique of science's self-understanding, he points out that Feyerabend's social and cultural critique suffers from a lack of engagement with relevant philosophical traditions, which offer more consistent criticism of scientism (2021: 176).

Kidd concurs with John Preston's assessment that Feyerabend's technology skepticism and his anti-scientist and anti-modernist views may be seen as an expression of a "reactionary romanticism", a desire to return to a prior state of culture where people were untroubled by intellectuals and their tendency to universalize concepts (Preston 2000: 621-622). Accusing Feyerabend of romanticism may be too strong. He does indeed repeatedly express appreciation for less scientifically influenced forms of life. But this should not be taken as the expression of a utopia that Western societies should turn back to a less modern state of development. Feyerabend seems to be eager to draw attention to the plurality of *possible* forms of life and the diversity and abundance of human cultures, without implying a hierarchy among cultures or a preference for a particular way of life. His criticism of modernity targets all forms of cultural chauvinism, particularly the chauvinism he identifies in some proponents of the scientist-rationalist worldview.

It goes without saying that Feyerabend's critique of scientific modernity must be interpreted against the cultural and political context of his time. From the late 1960s onwards, his writing was influenced by and reacted to contemporary cultural and political debates, including the student movements of the late 1960s, the emergence of environmentalist movements, and post-colonial discourses in the late 1970s and 1980s.[9] However, as Kidd (2021: 187-188) assesses, Feyerabend rarely engages substantively with contemporary philosophical contributions on topics like feminism or post-colonialism. He also does not systematically contextualize or relate his own cultural criticism to major works in 20th-century philosophy, such as Heidegger's *The Question Concerning Technology*, Husserl's *Crisis of the European Sciences,* or Adorno's and Horkheimer's *Dialectic of Enlightenment*.

This is rather unfortunate indeed, as it would be highly interesting to know to what extent Feyerabend was inspired by and how he integrated into his own

[9] Oberheim (2006: 24) describes Feyerabend's ability to constantly adapt to changing interest and attitudes as chameleon-like.

thinking approaches and concepts like Husserl's idea of the "lifeworld" (*Lebenswelt*), the notion that all knowledge must be grounded in lived experience. Feyerabend's lack of serious engagement with relevant contributions in the philosophical tradition that would seem to have a lot in common with his own criticism may reflect his disapproval of any sort of generalizing philosophical positions and "isms". His treatment of Husserl is telling in this respect. On the one hand, he calls *The Crisis of the European Sciences* "a remarkable essay"; on the other hand, he accuses Husserl of "astounding ignorance", "phenomenal conceit", and "sizeable contempt for anybody who lives and thinks along different lines" (1987: 274).

4. *A Feyerabendian ethos?*

It was probably John Preston who first identified voluntarism as "one of the most consistent themes in [Feyerabend's] work" (Preston 1997: 20). Indeed, Feyerabend's ethical-normative defense of pluralism is driven by the presupposition that scientific knowledge and the methodological rules guiding scientific knowledge production depend on our decisions. These decisions "can and must be evaluated by reference to our ideals" (Preston 1997: 21). Thus, science as well as the philosophy of science are based on and grounded in ethical values. Preston points out that Feyerabend's objections against theoretical monism must be read against the background of this strong voluntarism.[10]

More recently, Martin Kusch (2021) has discussed Feyerabend's voluntarism in relation to Bas van Fraassen's notion of a "philosophical stance". The basic idea of van Fraassen's stance concept is that philosophical accounts should not be understood as positions or doctrines but as stances. A stance can be characterized as a bundle of values, emotions, policies, and preferences. Kusch calls these bundles VEPPs (2021: 89). Like Preston, Kusch also concludes that Feyerabend is an epistemological voluntarist and that his philosophical account should be taken as a stance rather than a position or a doctrine.[11]

Feyerabend himself is very clear about the voluntarist nature of his approach to philosophy and, in fact, all knowledge. This aspect of Feyerabend's thinking has been present from a very early point. In "Knowledge Without Foundations", we find the following statement:

> It must again be repeated that we are here confronted with a real *decision*, that is, a real choice with a situation which has to be resolved on the basis of our demands and

[10] See also Preston's remarks on this in his Introduction to Feyerabend (1999).

[11] "Indeed, AM [*Against Method*] is adamant that pluralism is ultimately underwritten by humanitarianism, the system of VEPPs at the heart of epistemological anarchism" (Kusch 2021: 93).

preferences and which cannot be resolved by proof. […] [E]pistemology or the structure of the knowledge we accept, is grounded upon an ethical decision. This result is very different indeed from what seems to be the commonly accepted point of view. For it is usually assumed that the foundations of our knowledge are things which exist independently of human beings, which can be forgotten, misunderstood, overlooked but not eliminated with the help of a decision. This is quite correct provided we have already accepted a dogmatic point of view that works with certainties. Such a point of view will, of course, treat its own foundations as something that are given and cannot be influenced by human decisions (1961/1999: 72).

This passage helps to understand that Feyerabend's rejection of dogmatism, which he ascribes to monistic philosophies of science, results from his voluntarist stance. Every theory of science that aims at providing a foundation for scientific knowledge production, and which exists "independently of human beings" (i.e., of their decisions), thus becomes a target of Feyerabend's criticism. Foundational questions, according to Feyerabend, cannot be answered by merely studying the historical development of science (Kuhn) or by discovering the one and only method of scientific discovery (Popper). Foundational problems have to be *decided* in the literal sense of the word. The acceptance of a specific methodology or a set of methodological rules always implies a choice on the part of those who adopt those rules or propagate their acceptance. Because adopting a specific set of rules influences the kind and the content of the knowledge that emerges from it, and because that knowledge has implications for society and the lives of people, the decisions that go along with the choice of a methodology must be evaluated on ethical grounds. Such choices become a matter of moral responsibility. In Feyerabend's words:

The fact that almost any philosophical doctrine may find realization either in a *cosmology*, that is, in a theory of the universe that is capable of sensual representation, and/or in a *theory of man*, which may also be sensually realized in a corresponding society – this fact makes it very clear that the procedure leading to the adoption of a philosophical position cannot be *proof* (proof shows that no other position could possibly be realized), but must be a *decision* on the basis of preferences. […] [T]he problem responsible choice enters even the most abstract philosophical matters and that ethics is, therefore, the basis of everything else (1965: 219).

In a recent talk,[12] Matthew J. Brown has discussed this voluntarist aspect of Feyerabend thinking from a pragmatist-axiological perspective. Brown points out that, for Feyerabend, the aim of science is to promote human flourishing,

[12] At the 17th International Congress on Logic, Methodology and Philosophy of Science and Technology CLMPS, University of Buenos Aires, July 2023.

and that the progress of science cannot be judged in a purely epistemic way. Therefore, to foster the promotion of human flourishing, Feyerabend engages in a historical and philosophical critique of the epistemological obstacles to that flourishing. Knowledge producing activities are thus constrained and guided by ethical and moral values.

Feyerabend seems to be using the term "humanitarianism" as a placeholder for values that he thinks should constrain science. However, it remains unclear from Feyerabend's writings what he exactly means by "humanitarianism". Feyerabend at no point argues in favor of any general principles. Neither does he associate himself with an ethical system or engage in any sort of meta-ethical argumentation. It is in fact not at all surprising that we do not find a systematically developed ethical position in Feyerabend's work. In his autobiography, he clearly states his aversion against any sort of overly generalizing philosophical accounts:

> So there were now two types of tumors to be removed – philosophy of science and general philosophy (ethics, epistemology, etc.) – and two areas of human activity that could survive without them – science and common sense (1995: 142).

The question thus remains: How is Feyerabend's self-ascribed humanitarianism to be understood and what are the values that characterize his humanitarian outlook? Preston (1997: 21) lists humanity, respect for the individual, happiness, joy, pleasure, imagination, sense of humor, and capriciousness as goals that Feyerabend frequently appeals to. Kusch (2021) focuses on the VEPPs of Feyerabend's humanitarianism. However, Kusch does not describe in detail what Feyerabend's VEPPs exactly are.

If we were to characterize Feyerabend's ethical-normative stance in any way, it might best be understood as a form of moral particularism that denies the usefulness of general moral principles or meta-ethical theories and emphasizes the importance of context for moral judgments. It is thus futile to look for general moral or meta-ethical principles in Feyerabend's works. What is possible, though, is to extract a set of attitudes that appear frequently in Feyerabend's writing. These include:

- A strong anti-dogmatism, anti-authoritarianism, and anti-hegemonism (theoretical, cultural, political).[13]
- A strong emphasis on the cultivation of individuality and the free development of individuals.

[13] In his presentation at a Feyerabend workshop at the University Mohammed VI Polytechnic in February 2024, Matteo Collodel identified anti-authoritarianism as a constant underlying motive of Feyerabend's thinking that must be interpreted in light of his traumatic coming of age in Nazi-occupied Austria and the ensuing experiences during World War II.

- A strong emphasis on the free development of group identities, cultures, and "traditions".
- Tolerance of heterodox opinions and practices.
- Openness towards non-rationalistic narratives (myths, stories, spiritual & religious traditions).
- Anti-rationalism insofar as the scientistic-rationalistic worldview is taken as just one tradition among many.
- Intellectual eclecticism and universalism manifesting itself in the belief in the unity of science and art as well as the unity of natural science and humanities.

To be sure, this list is probably incomplete as it would be easy to ascribe more or different attitudes to Feyerabend. It also cannot be ruled out that upon closer analysis, one would be able to identify tensions and contradictions in these attitudes (such as, for example, tensions between progressive and conservative or between individualist and collectivist attitudes).

A focal point of Feyerabend's ethical stance seems to be his firm belief in a common human nature and a shared humanity. His normative defence of pluralism seems to be driven by the assumption that the bonds created by our universally shared humanity are stronger than the divisions emerging from different philosophies, traditions, cultures, ideologies, systems of belief, or ways of life. This humanistic universalism is expressed most clearly in the somewhat obscure statement that potentially every culture is all cultures:

> If every culture is potentially all cultures, then cultural differences lose their ineffability and become *special and changeable manifestations of a common human nature*. Authentic murder, torture, and suppression become ordinary murder, torture, and suppression, and should be treated as such (1994: 21).

This passage may best be interpreted as the view that the abundance and diversity of human cultures are ultimately manifestations of a common human nature and that actions that inflict harm on humans are morally reprehensible independently of the cultural or ideological reasons that might justify such actions. If anything, it is the belief in a shared human nature and the protection and flourishing of individuals and traditions that constitute the central elements of Feyerabend's humanitarianism.

5. *The relevance of Feyerabend's pluralism for science-based policy*

The fact that the ethical-normative foundation of Feyerabend's pluralism is systematically weak and not well-developed does not render his pluralistic account futile. In a series of papers, Simon Lohse and I have used Feyerabend's

pluralism as a philosophical point of reference to discuss epistemic shortcomings in the science-based policy advice during the COVID-19 pandemic (Lohse *et al.* 2020; Bschir *et al.* 2022; see also Bschir *et al.* 2024). Our goal was to show that Feyerabend's arguments in favor of epistemic pluralism are useful to highlight the importance of inner-scientific pluralism, as well as the inclusion, of extra-scientific perspectives in science-based decision-making processes.

The scientifically informed policy responses to the COVID-19 pandemic were subject to a variety of criticisms. One type of criticism targeted methodological aspects of the science that informed policies during the pandemic. There were controversies surrounding the quality of the available evidence on which policy measures were based (e.g. Ioannidis 2020). Others pointed at the uncertainties of model-based predictions (e.g. Saltelli *et al.* 2020; Chin *et al.* 2021), and doubts were raised about the transferability of projections and policy interventions between different countries (e.g. Sebhatu *et al.* 2020). Another important line of criticism highlighted an imbalance in favor of biomedical and epidemiological perspectives at the expense of others, in particular of socio-economic (e.g. Broadbent *et al.* 2020) and societal perspectives (e.g. Lohse *et al.* 2021). The latter kind of criticism may be interpreted as highlighting a lack of epistemic pluralism in the scientific knowledge that informed policies during the pandemic.

As seen in the previous sections, Feyerabend clearly emphasizes the importance of pluralism within science. Pluralism is epistemologically beneficial as it fosters the detection of errors and limitations in received approaches. But, according to Feyerabend, pluralism also has implications for the relationship between science and society and the role of scientific expertise. Since he does not limit the domain of permissible alternatives to scientific theories, his account also provides good arguments for the inclusion of extra-scientific perspectives. In order to bring Feyerabend's pluralism to bear in the analysis of the epistemic shortcomings in the response to the COVID-19 pandemic, we reconstruct Feyerabend's pluralist account along four key elements (see Bschir *et al.* 2022: 440-441).

The first element is fallibilism. Feyerabend's epistemic pluralism is based on the supposition that it is reasonable to assume the fallibility of even the most successful scientific theories and methods.

The second element is the emphasis on the importance of alternatives in revealing the shortcomings in established methods and theories. Alternatives are necessary, according to Feyerabend, to unearth problems and limitations, which would remain invisible in the absence of the alternatives.

The third element, proliferation, is closely related to the second. Because alternatives can highlight problems and limitations of any given theory or method, the active development and simultaneous application of a variety of ap-

proaches can help to avoid myopia with respect to a given epistemic problem.

The fourth element is the extension of pluralism beyond scientific approaches. For Feyerabend, non-scientific forms of knowledge (practical knowledge or the knowledge of local communities) can constitute important correctives to scientific approaches, as they can provide insights into aspects of a given problem (i.e. the question of how to manage a global pandemic at the local level) that would remain invisible in a strictly scientific perspective. This is relevant for policy making, because the insights that such non-scientific perspectives can provide, may be highly relevant to those affected by policy measures.

Using this reconstruction of Feyerabend's epistemic pluralism, we were able to highlight the importance of pluralism in the context of COVID-19 policies on three levels.

Intra-disciplinary pluralism: Within a discipline (e.g. epidemiology) perusing a multitude of methodologies and theoretical approaches can help to highlight the shortcomings of each alternative. This is particularly important in situations where scientific results can have a direct influence on policy. The application of a variety of epidemiological modeling approaches (e.g., standard SIR models alongside network and agent-based models) to inform policy during the pandemic is a good case in point.

Interdisciplinary pluralism: Although the reconstruction of Feyerabend's pluralism in Section 2 against the background of his critique of Kuhn mainly focuses on pluralism within a discipline, we argue that it also provides good arguments for interdisciplinary pluralism. In the context of the COVID-19 pandemic, this would have meant avoiding a concentration on biomedical and epidemiological perspectives alone, and including insights from other disciplines and the social sciences. The latter are particularly important not only for assessing the societal consequences of policy measures (e.g., school closures or lockdowns of elderly homes) but also for identifying social effects of policies in different sections of society (see e.g., Sari *et al.* 2021).

Science-transcending epistemic pluralism: Finally, we can use Feyerabend's account to support arguments in favor of science-transcending epistemic pluralism, highlighting the importance of extra-scientific perspectives and local forms of knowledge, for example, for the detection of unintended negative consequences or in the design, implementation, and assessment of policy measures.

It is beyond the scope of this summary to discuss the many practical challenges that the implementation of pluralism in science-based policy-making

poses (for a discussion of these challenges see Section 5 in Bschir *et al.* 2022 and Bschir *et al.* 2024). While some of these challenges are serious and need to be addressed in future studies, our goal was to provide good arguments for more pluralism at the science-policy interface based on Feyerabend's account, thus highlighting the potential relevance of Feyerabend's pluralist stance in current contexts. Considering Feyerabend's voluntarism, it also becomes clear that implementing more pluralism at the science-policy interface is not only epistemologically desirable, but that it implies a deliberate choice that can and must be evaluated in light of its practical consequences and ethical implications.

6. *Conclusion*

I have examined the hitherto underexplored relationship between Feyerabend's pluralist philosophy of science and the ethical-normative humanitarian account that underlies his epistemological position. I have shown that while his epistemic pluralism is well-argued and systematically developed, his ethical position is systematically rather weak. The main reason for this is that Feyerabend explicitly denied the validity and legitimacy of any sort of general philosophical or ethical accounts. His approach to ethics can best be characterized as a form of moral particularism that does not provide any sort of general ethical principles but, at best, puts forward a set of implicit values and attitudes.

The systematic weakness of Feyerabend's ethical-normative account does not diminish the relevance of his pluralistic philosophy of science. Quite on the contrary, the latter can be brought to bear in the current contexts, for example in science-based policymaking during the COVID-19 pandemic. Feyerabend's pluralism is useful to address the epistemic shortcomings in the policy responses to the pandemic, such as the imbalance in favor of biomedical and epidemiological perspectives at the expense of socioeconomic and societal perspectives.

Feyerabend's account of pluralism reconstructed along four key elements – fallibilism, the importance of alternatives, proliferation, and the extension beyond scientific approaches – provides robust arguments for intra-disciplinary, interdisciplinary, and science-transcending epistemic pluralism. His arguments support the inclusion of diverse methodologies and theoretical approaches within disciplines, the integration of insights from different disciplines, as well as the consideration of non-scientific forms of knowledge in policymaking.

Implementing pluralism in science-based policymaking is not only epistemologically desirable but also involves deliberate ethical choices that must be evaluated in light of their practical implications. It remains a matter of future research to investigate the potential relevance of Feyerabend's epis-

temic pluralism in other contexts, such as, for example, economic or environmental policy.

Karim Bschir
University of St. Gallen
karim.bschir@unisg.ch

Acknowledgments: I would like to thank the organizers of the workshop Pierluigi Barrotta, Gustavo Cevolani, Eleonora Montuschi, Roberto Gronda, and Luca Tambolo for the invitation. Conversations with Grazia Borrini-Feyerabend, Matthew Brown, Hasok Chang, Matteo Collodel, Steven Fuller, Michael Hampe, Paul Hoyningen-Huene, Eric Oberheim, and John Preston have provided me with helpful insights on the relationship between Feyerabend's pluralism and humanitarianism. I also thank Simon Lohse and two anonymous referees for constructive feedback on an early draft.

References

Broadbent, Alex, Benjamin T.H. Smart, 2020, "Why a One-Size-Fits-All Approach to COVID-19 Could Have Lethal Consequences", in *The Conversation*, accessed in February 2024, http://theconversation.com/why-a-one-size-fits-all-approachto-covid-19-could-have-lethal-consequences-134252.

Bschir, Karim, 2015, "Feyerabend and Popper on Theory Proliferation and Anomaly Import: On the Compatibility of Theoretical Pluralism and Critical Rationalism", in *HOPOS: The Journal of the International Society for the History of Philosophy of Science*, 5, 1: 24-55.

—, Simon Lohse, 2022, "Pandemics, Policy, and Pluralism: A Feyerabend-inspired Perspective on COVID-19", in *Synthese*, 200, 6: 441.

—, Simon Lohse, 2024, "Taking Pluralism Seriously: A New Perspective on Evidence-Based Policy", in *Science and Public Policy*, 51, 3: 553-556.

Collodel, Matteo, 2016, "Was Feyerabend a Popperian? Methodological issues in the History of the Philosophy of Science", in *Studies in History and Philosophy of Science Part A*, 57: 27-56.

Farrell, Robert P., 2003, *Feyerabend and Scientific Values: Tightrope-Walking Rationality*, Kluwer Academic Publishers, Dordrecht.

Feyerabend, Paul Karl, 1961/1999, "Knowledge without Foundations", as reprinted in Feyerabend 1999: 50-77.

—, 1963/1999, "How to be a Good Empiricist: A Plea for Tolerance in Matters Epistemological", as reprinted in Feyerabend 1999: 78-103.

—, 1965, "Problems of Empiricism", in Colodny, Robert G., ed. *Beyond the Edge of Certainty: Essays in Contemporary Science and Philosophy*, Prentice-Hall, Englewood Cliffs: 145-260.

—, 1970, "Consolations for the Specialist", in Lakatos, Imre, Alan Musgrave, eds., *Criticism and the Growth of Knowledge: Volume 4: Proceedings of the International Colloquium in the Philosophy of Science, London 1965*, Cambridge University Press, Cambridge: 197-230.

—, 1975a, "How to Defend Society Against Science", in *Radical Philosophy*, 11: 3-8.

—, 1975b, *Against Method: Outline of an Anarchistic Theory of Knowledge*, New Left Books, London.

—, 1976, "On the Critique of Scientific Reason", in Howson, Colin, ed., *Method and Appraisal in the Physical Sciences: The Critical Background to Modern Science 1800-1905*, Cambridge University Press, Cambridge: 309-339.

—, 1978, *Science in a Free Society*, Verso, London and New York.

—, 1980, "Democracy, Elitism, and Scientific Method", in *Inquiry*, 23, 1: 3-18.

—, 1985, "Autoritäre Konstruktionen oder demokratische Entschlüsse? Bemerkungen zum Problem des kulturellen Pluralismus", in Schmied-Kowarzik, Wolfdietrich, ed., *Objektivationen des Geistigen: Beiträge Zur Kulturphilosophie in Gedenken an Walther Schmied-Kowarzik*, Dietrich Reimer, Berlin: 137-144.

—, 1987, *Farewell to Reason*, Verso, London and New York.

—, 1993, *Against Method*, 3rd ed., Verso, London.

—, 1995, *Killing Time: The Autobiography of Paul Feyerabend*, The University of Chicago Press, Chicago.

—, 1999, *Knowledge, Science and Relativism: Philosophical Papers, Volume 3*, edited by John Preston, Cambridge University Press, Cambridge.

—, 2011, *The Tyranny of Science*, edited by Eric Oberheim, Polity Press, Cambridge.

Gillies, Donald, 2024, "Feyerabend's Criticisms of Kuhn", in *Filozoficzne Aspekty Genezy*, 20: 2.

Hoyningen-Huene, Paul, 1995, "Two Letters of Paul Feyerabend to Thomas S. Kuhn on a Draft of the Structure of Scientific Revolutions", in *Studies in History and Philosophy of Science*, 26: 353-387.

—, 2000, "Paul Feyerabend and Thomas Kuhn", in Preston, John, Gonzalo Munévar, David Lamb, eds., *The Worst Enemy of Science? Essays in Memory of Paul Feyerabend*, Oxford University Press, New York.

—, 2006, "More Letters by Paul Feyerabend to Thomas S. Kuhn on *Proto-Structure*", in *Studies in History and Philosophy of Science*, 37: 610-632.

Ioannidis, John P.A., 2020, "In the Coronavirus Pandemic, We're Making Decisions Without Reliable Data", *STAT*, accessed in February 2024, https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronaviruspandemic-takes-hold-we-are-making-decisions-without-reliable-data.

Kidd, Ian James, 2021, "Feyerabend, Science and Scientism", in Bschir, Karim, Jamie Shaw, eds., *Interpreting Feyerabend. Critical Essays*, Cambridge University Press, Cambridge: 172-190.

Kuby, Daniel, 2021, "Feyerabend's Re-Evaluation of Scientific Practice: Quantum Me-

chanics, Realism and Niels Bohr", in Bschir, Karim, Jamie Shaw, eds., *Interpreting Feyerabend. Critical Essays*, Cambridge University Press, Cambridge: 132-156.

Kuhn, Thomas S., 1962, *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago.

Kusch, Martin, 2021, "Epistemological Anarchism Meets Epistemic Voluntarism: Feyerabend's *Against Method* and van Fraassen's *The Empirical Stance*", in Bschir, Karim, Jamie Shaw, eds., *Interpreting Feyerabend. Critical Essays*, Cambridge University Press, Cambridge: 89-113.

Lloyd, Elisabeth A., 1997, "Feyerabend, Mill, and Pluralism", in *Philosophy of Science*, 64: S396-S407.

Lohse, Simon, Karim Bschir, 2020, "The COVID-19 Pandemic: A Case for Epistemic Pluralism in Public Health Policy", in *History and Philosophy of the Life Sciences*, 42: 58.

—, Stefano, Canali, 2021, "Follow *the* Science? On the Marginal Role of the Social Sciences in the COVID-19 Pandemic", in *European Journal for Philosophy of Science*, 11: 99.

Martin, Michael, 1972, "Theoretical Pluralism", in *Philosophia*, 2: 341-350.

Oberheim, Eric, 2006, *Feyerabend's Philosophy*, de Gruyter, Berlin.

Preston, John, 1997, *Feyerabend: Philosophy, Science and Society*, Polity Press, Cambridge.

—, 2000, "Review of *Conquest of Abundance: A Tale of Abstraction versus the Richness of Being* By Paul Feyerabend edited by Bert Terpstra University of Chicago Press", in *Philosophy*, 75: 613-626.

Saltelli, Andrea, Gabriele Bammer, Isabelle Bruno, Erica Charters, Monica Di Fiore, Emmanuel Didier, Wendy Nelson Espeland, John Kay, Samuele Lo Piano, Deborah Mayo, Roger Pielke Jr, Tommaso Portaluri, Theodore M. Porter, Arnald Puy, Ismael Rafols, Jerome R. Ravetz, Erik Reinert, Daniel Sarewitz, Philip B. Stark, Andrew Stirling, Jeroen van der Sluijs, Paolo Vineis, 2020, "Five Ways to Ensure that Models Serve Society: A Manifesto", in *Nature*, 582 (7813): 482-484.

Sari, Elif, Felix Bittmann, Christoph Homuth, 2021, "Explaining Inequalities of Homeschooling in Germany During the First COVID-19 Lockdown", in *SocArXiv Papers*.

Sebhatu, Abiel, Karl Wennberg, Stefan Arora-Jonsson, Staffan I. Lindberg, 2020, "Explaining the Homogeneous Diffusion of COVID-19 Nonpharmaceutical Interventions Across Heterogeneous Countries", in *Proceedings of the National Academy of Sciences*, 117: 21201-21208.

Sirtes, Daniel, Eric Oberheim, 2006, "Einstein, Entropy, and Anomalies", in Alimi, Jean-Michel, André Füzfa, eds., *Albert Einstein Century International Conference: American Institute for Physics Conference Proceedings*, 861: 1147–1154.

Tambolo, Luca, 2015, "A Tale of Three Theories: Feyerabend and Popper on Progress and the Aim of Science", in *Studies in History and Philosophy of Science Part A*, 51: 33-41.

van Strien, Marij, 2020, "Pluralism and Anarchism in Quantum Physics: Paul Feyerabend's Writings on Quantum Physics in Relation to his General Philosophy of Science", in *Studies in History and Philosophy of Science Part A*, 80: 72-81.

Feyerabendian pluralism in practice: Lessons from the Di Bella case

Luca Tambolo, Gustavo Cevolani

Abstract: This paper contrasts two ingredients of Feyerabendian pluralism: the idea that the proliferation of theories and methods is good for science (the "limited pluralism" view) and the view of knowledge as an ever increasing ocean of mutually incompatible alternatives (the "oceanic" view). We argue that, in order for Feyerabendian pluralism to be tenable, the limited pluralism view should be decoupled from the oceanic one, and the latter rejected. We use as a case study that of Luigi Di Bella, an obscure Italian physician who in 1997-1998 suddenly became a national celebrity as the self-proclaimed discoverer of "the cure for cancer". When the case erupted, no evidence of the efficacy of Di Bella's unconventional approach to cancer treatment was available, and the relevant experts concurred that the so-called "Di Bella method" (DBM) did not show any promise. Yet, the Parliament passed a piece of ad hoc legislation authorizing a series of phase II state-funded clinical trials aimed at assessing the DBM. Asking what course of action a Feyerabendian pluralist would have recommended in this scenario allows one to probe into the – limited, as it turns out – validity of some of Feyerabend's views on theoretical pluralism.

Keywords: Feyerabend, Theoretical pluralism, Limited pluralism, "Oceanic" pluralism, Di Bella case.

1. *Introduction*

Feyerabend passionately championed a strong pluralism concerning science, its theories, and its methods. In doing so, he defended both the view that genuine knowledge requires the proliferation of alternatives to the dominant theory, even when the latter enjoys an impressively high degree of support (to which we refer as the "limited pluralism" view), and the view of knowledge as an ever increasing ocean of mutually incompatible alternatives (the "oceanic" view). Both of these views provided important arguments for Feyerabend's no less passionate criticism of experts who, by virtue of their alleged excellence in some domain of scientific inquiry, demand a privileged role in democratic policy-making – a demand against which he railed especially in *Science in a Free Society* (1978). It is an open question whether and to what degree appropriately modified versions of Feyerabend's views on the importance of pluralism and

philinq XII, 1-2024, pp. 145-161
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.556
Submitted: 4 June 2024
Accepted: 17 August 2024

the proper role of experts in public decision-making can be combined within a satisfactory account of expert policy advice (see, e.g., Brown 2021; Bschir *et al.* 2022; Shaw 2021). With a view to contributing to the ongoing discussion on such an issue, in this paper we focus on a historical episode that provides one with excellent material to probe into the validity of some of Feyerabend's views on theoretical pluralism: the so-called "Di Bella case".

In a nutshell, Luigi Di Bella was an obscure Italian physician who, at the end of 1997, hit the media headlines as the self-proclaimed discoverer of no less than the cure for cancer. A harsh critic of approved cancer treatments, Di Bella fueled a national controversy concerning the respective merits of chemotherapy, to which he disparagingly referred as the product of "official oncology", and of his unconventional therapeutic approach, for which he was unable to offer evidence of safety and efficacy. Although very heated, the controversy was short-lived, especially in comparison to the well-known case of laetrile (see Markle *et al.* 1980): already by late 1998 the support of media outlets that at the end of the previous year had made Di Bella an instant celebrity was a thing of the past, and Di Bella quickly disappeared from public view. A most remarkable feature of the case was that in response to the widespread popular support gained by Di Bella, at the beginning of 1998 the Parliament authorized – against the advice of the relevant community of scientific and medical experts – a series of clinical trials aimed at testing Di Bella's treatment.

In such a scenario, we argue, Feyerabendian pluralists could easily have found themselves on the wrong side of the fence separating Di Bella's supporters from his critics. More specifically, we show that Feyerabendian pluralists could have found themselves in agreement with the troublesome decision to authorize the trials of the "Di Bella Method" (DBM, for short) – namely, experiments on human subjects conducted in spite of the fact that there was no reason to presume the DBM's efficacy. As we argue, this suggests that in order for Feyerabendian pluralism to deliver the goods that Feyerabend quite correctly claimed it can help us to achieve, one should decouple the limited pluralism view from the oceanic view of knowledge, espoused in Feyerabend's writings at least since the mid-1960s (e.g., 1965/1981; 1975; 1993).

We proceed as follows. In Section 2 we summarize the facts of the case, emphasizing in particular the status of radical alternative to the dominant theory that Di Bella claimed for his approach to the treatment of tumors. In Section 3 we show that based on the limited pluralism view, revolving around Feyerabend's pluralistic model of theory testing, Feyerabendian pluralists would have disagreed with the decision to authorize the trials. However, the oceanic view requires one to drop any condition of factual adequacy of the alternatives to the dominant theory that one welcomes to enter into a scientific discussion.

As we show in Section 4, based on the oceanic view, Feyerabendian pluralists would have agreed with the authorization of the trials. This, we argue, speaks strongly against such view. In Section 5 we discuss a possible rebuttal of our analysis on the part of Feyerabendian pluralists, based on the distinction drawn by Feyerabend between "respectable thinkers" and cranks. The distinction is of course pretty sensible but, as we argue, it does not impinge on our claim that the oceanic view should be dropped. In Section 6 we offer some brief concluding remarks.

2. *A crash course in the Di Bella case*

Luigi Di Bella (1912–2003) was 85 years old when he became a national celebrity. A medical doctor and retired professor of Physiology at the University of Modena, until late 1997 Di Bella was unknown to the general public. His rise to fame, as sudden as unexpected, was due to his proclaimed ability to treat cancer with the DBM – the "Di Bella Method" or "Di Bella Multitherapy" – that he himself had invented and claimed to have perfected over the decades on the patients of his private practice.

Di Bella advertised the DBM as an approach to the treatment of tumors exhibiting much higher levels of safety and efficacy than the current standards of care, especially chemotherapy – which he harshly criticized. In fact, in Di Bella's own words, the DBM revolves around the idea that, unlike in chemotherapy, one should aim at "promoting the biological conditions conducive to preventing further proliferation of neoplastic cells without destroying the pre-existing ones" (Di Bella *et al.* 2012: 248). An appropriate combination of substances, administered by the treating physician based on their clinical judgment concerning the situation of the individual patient, is supposed to block the production of the growth hormone, which Di Bella viewed as one of the most important factors responsible for the formation of neoplastic cells. In the favorable conditions created by the administered substances – to which many referred as the "Di Bella cocktail" – "the destruction of neoplastic cells takes place thanks to the competition instituted between the healthy cell, which grows, and the inability of the neoplastic cell to exploit the resources available" (Di Bella *et al.* 2012: 248).[1]

In 1997 the relevant scientific and medical communities were unaware of the DBM, since no body of scientific literature described it or suggested,

[1] This summary of the basics of the DBM was offered by Di Bella during a public event in Parma on April 4, 1998. Di Bella's words have been very aptly chosen and translated into English within the piece from which we quote, written for his centennial and co-authored by his sons Giuseppe and Adolfo. For a systematic presentation of the DBM see, e.g., Di Bella (2010).

let alone documented, its alleged safety and efficacy (a situation that hasn't changed in the meantime).[2] But Di Bella had by then built for himself a solid if underground reputation, within a following of devoted patients, as a specialist in the treatment of tumors. And the DBM became the subject of a heated national controversy after groups of organized patients staged public demonstrations calling for the Italian government to include somatostatin and ocreotide – both ingredients of Di Bella's cocktail – in the list of drugs provided free of charge for the compassionate treatment of patients with all types of tumor (in that moment, somatostatin and ocreotide were provided for free only for selected indications, for which evidence was available). Had the request been granted, the Italian National Health Service would have been bound to provide somatostatin and ocreotide for free to all cancer patients, thereby resolving the demonstrators' main discontent: because of the high cost of somatostatin, which in public discourse became the "flagship drug" of the DBM, treatment with the DBM was in fact inaccessible to many who would have wanted it. For the government, however, accepting the request would have amounted to overruling a number of decisions, roundly criticized by the demonstrators, issued based on the available evidence by the National Drug Committee between 1996 and 1997.

The national media jumped on the story in mid-November 1997, and covered it in real time in countless TV and radio news and broadcasts, as well as in literally thousands of newspapers and magazine articles. More or less at the same time, various local judges started to order health authorities in different parts of the country to provide somatostatin for free to cancer patients requesting it. And the now suddenly widespread conviction that the safety and efficacy of the DBM was a matter of course received support also from some opposition parties, wishing to use popular discontent with the difficulties faced by patients who could not afford the costs of the DBM as a means to attack the government. All of this happened in spite of the fact that in November 1997, there was no publicly available evidence suggesting the efficacy of the DBM (or ruling out its toxicity, for that matter), only Di Bella's and his followers' claims. Yet, much to the worry and dismay of scientists and clinicians, an unknown but sizable number of patients abandoned approved treatments in favor of the DBM. In light of such a development, of the fact that Italy had become a country divided between supporters and critics of Di Bella and his method, and of the request from many quarters to subject Di Bella's claims to proper

[2] To mention but one instance, Ernst (2021: 113) includes the DBM in his list of "so-called alternative medicine" (SCAM) for cancer and comments: "Even though Di Bella published several further investigations, there is still no sound evidence that this treatment is effective in any type of cancer". The "further investigations" to which Ernst refers, published after Luigi Di Bella's death in 2003, are mainly due to Giuseppe Di Bella.

scientific testing, the government quickly, albeit reluctantly opened the door to the possibility of a state-funded experimentation of the DBM.

To cut a convoluted story short, in February 1998 the Parliament passed a piece of legislation authorizing a series of phase II clinical trials – which started in March of the same year – aimed at testing the safety and efficacy of the DBM. The results of the trials, published in the *The British Medical Journal* in 1999, infuriated champions of the DBM, which the experimenters deemed not promising enough, in terms of its efficacy, to warrant further clinical testing (Italian Study Group for the Di Bella Multitherapy Trials 1999). To this day, the circle of Di Bella's heirs and supporters claims that the trials were a fraud orchestrated by the scientific and pharmaceutical establishment owing its prestige and wealth to the dominance of chemotherapy, and therefore hostile to the DBM, and denounces flaws both in the design and in the execution of the experimental work (see, e.g., Di Bella 2012: 391-408). Di Bella himself openly voiced his complete distrust of the experimenters even before the trials took place. For instance, when his first biographer asked him whether he thought that the trials would warrant a fair assessment of the DBM, he replied by noting that "the field of oncology is dominated by chemotherapists. They are natural enemies of my method, since it radically questions theirs" (Di Bella *et al.* 1998: 8). In any case, after some preliminary results of the trials were announced by mid-1998, popular support for the DBM quickly waned, and by late 1998 Di Bella had lost the prominence on the public scene that he had gained only a year earlier.

What matters for our present purposes, though, is the very fact that the trials were performed. As already mentioned, in 1997 there was no publicly available evidence suggesting the safety and efficacy of the DBM. And of course, extant legislation provided that before starting clinical experimentation of any drug or treatment, one needed appropriate evidence in hand, indicating at least its safety and possible promise. Thus, in December 1997 Minister of Health Rosy Bindi ordered Di Bella to turn in a sample of the records of patients which he claimed to have treated with the DBM over the decades, contained in his archive, in order to ascertain whether the conditions for authorizing the trials obtained. Di Bella begrudgingly complied and handed in a sample consisting of less than one hundred items.

In the view of the experts who examined the records of Di Bella's patients on behalf of the Ministry of Health, auditioned by the relevant Committee of the Chamber of Deputies in January 1998, the conditions for authorizing the trials did not obtain,[3] since nothing indicated the possible efficacy of the DBM

[3] See, e.g., the remarks by members of the National Cancer Advisory Committee and of the National Drug Committee in (Minutes of the January 15, 1998, Parliamentary Hearing: 52 and 60 respectively).

– which Di Bella on numerous occasions advertised as effective in about the 90 per cent of the cases. In spite of this, the government ended up supporting the decision to subject Di Bella's claims to experimental scrutiny. Bindi, auditioned by the same Committee a few days after the Ministry's experts, pointed out that of course, the path followed to get at the eve of the approval of the trials was "in a certain sense anomalous" (Minutes of the January 20, 1998 Parliamentary Hearing: 86). In fact, no drug manufacturing company had requested a trial for any of its products, nor had Di Bella demanded an experimentation of the DBM. It was in response to public pressure on the government, Bindi emphasized, that the Ministry of Health had started to search for reasons to justify the trials. And in Bindi's view – plainly contradicting that of the experts – such reasons had been found after all. Nevertheless, Bindi remarked, the authorization of the trials had of course much to do with "what we can by now define a social issue" (Minutes of the January 20, 1998, Parliamentary Hearing: 88), namely, that of cancer patients abandoning approved treatments in favor of the DBM.

Admittedly far from providing the reader with an exhaustive reconstruction of the episode, our summary sets the scene for asking how a Feyerabendian pluralist could have dealt with the Di Bella case. More specifically, we wish to ask whether a Feyerabendian pluralist would have been in favor or against the authorization of the trials of the DBM.

3. *The limited pluralism view vis-à-vis the Di Bella case: do not authorize the trials*

Scholars have long emphasized that Feyerabend's occasional defenses of unorthodox positions and doctrines should not be taken at face value (Lloyd 1997/2000).[4] Yet, there can be no doubt that Feyerabend vociferously defended pluralism throughout all of his career. In a number of landmark essays published in the early 1960s, he put forward the principle of proliferation: a good empiricist should invent and elaborate alternatives to the theory that happens to be the dominant one at any point in time, no matter how favorably the relevant scientific community views the dominant theory (see, e.g., 1963/1999a; 1965/1981). Towards the end of the 1960s, Feyerabend came to view the proliferation not only of theories, but also of the methods used by scientists to appraise them as necessary for pluralism. In *Against Method* (1975), he therefore supplemented his earlier plea for theory proliferation with the

[4] To mention but one instance, the precise significance, as well as the merit, of Feyerabend's infamous advocacy of astrology in *Science in a Free Society* (1978: 91-96) has been the subject of quite some controversy in recent years (Kidd 2016a, 2016b; Pigliucci 2016a, 2016b, 2016c; Shaw 2017).

defense of methodological pluralism – that is, the proliferation of methods or, as he provocatively put it, anarchism. Moreover, Feyerabend's pluralism extended well beyond the domain of the methodology of scientific inquiry: starting from the 1970s, he forcefully argued that various non-scientific ways of investigation provide their practitioners with genuinely valuable insights into the workings of nature (see, e.g., 1987). And what matters the most for our present purposes, the DBM exhibited some of the features that Feyerabendian pluralists are bound to find attractive in an alternative to a dominant theory.

For a start, Di Bella himself described the DBM as a radical alternative to chemotherapy: while the latter targets neoplastic cells, the former aims at the creation of a situation in which healthy cells proliferate and manage, so to speak, to "overcome" neoplastic ones. Second, Di Bella's therapeutic approach revolved around the alleged identification of the mechanism responsible for the proliferation of neoplastic cells. In fact, as mentioned above, Di Bella viewed the production of the growth hormone as one of the main causes of cancer – a causal claim not supported by the evidence available back in 1997 (or by the evidence emerged since then). Sure, the DBM did borrow elements from current cancer treatments. For instance, synthetic analogues to somatostatin had been approved for the treatment of so-called "apudomas", a group of heterogeneous tumors arising from the diffuse neuroendocrine system. Di Bella, however, recommended somatostatin (together with the synthetic analog ocreotide and the other substances included in his cocktail) for the treatment of *all* tumors, not just of apudomas – a recommendation based on his account of the mechanism that causes neoplastic cells to proliferate. Moreover, to this day Di Bella's followers claim that the "official" statistics documenting the efficacy of chemotherapy are simply a fraud, and that chemotherapy does not cure, but rather kills patients (Di Bella 2012: 391-408).

Of course, it is precisely the lack of a highly effective treatment for all kinds of tumor that creates the space for the emergence of what Ernst (2021) dubs "so-called alternative medicine (SCAM) for cancer". In the case of Di Bella, the alternative to the dominant if largely imperfect therapeutic approach was not only theoretical, but also methodological. In fact, while emphatically insisting on his own scientific credentials – he claimed to have been shortlisted for the Nobel Prize in Medicine – Di Bella rejected methods for the appraisal of drugs and treatments that, post-World War II, had become a staple of a scientific approach to medicine. In particular, Di Bella consistently denied that statistics and randomized controlled trials are methods that allow one to properly assess the safety and efficacy of a drug or treatment.

Auditioned by the Committee for Social Affairs of the Chamber of Deputies in January 1998, Di Bella claimed that, "based on a rough calculation", the

material contained in his archive of record of patients was "more than enough to confirm what I am saying, that is, the efficacy of this treatment" (Minutes of the January 14, 1998, Parliamentary Hearing: 19). Indeed, Di Bella lost no occasion to brag about the successes achieved by the DBM, which he declared effective in 90 per cent of the cases, and he and his sidekicks insisted that at least ten thousand patients had undergone treatment with the DBM. Concerning the possibility of a statistical elaboration of the information contained in the medical records kept in his archive – which many demanded, as one of the means to assess whether or not the DBM was a promising approach – Di Bella agreed that the task, in which he himself had not engaged, was certainly feasible. He however hastened to add: "for what statistics is worth, given the enormous variability of the substrate on which the very therapy has acted" (Minutes of the January 14, 1998 Parliamentary Hearing: 19). In other words, Di Bella viewed such tools as protocols, statistical analyses, controlled studies, etc., as utterly inadequate to the purpose of assessing the DBM. The son of a bygone era of medicine, Di Bella regarded current methods to scientifically test drugs and treatments as either worthless or, worse, potentially misleading, since in his view, the clinical eye is all that really matters: the oncologist "ought to be an internist among the most endowed one can imagine, because he has to be able to unveil and interpret what happens in the organism of the patient" (Di Bella *et al*. 2012: 248).

In short, it is safe to say that Di Bella defied what he disparagingly called "orthodoxy" both at the theoretical and at the methodological level. The question then arises of how Feyerabedian pluralists would have dealt with the case. Would the embrace of theoretical and methodological pluralism, as Feyerabend understood them, have led to recommend the authorization of the trials?

Let us start pondering this question by recalling that in the early 1960s, Feyerabend made a name for himself as the critic of a set of ideas concerning theory testing which he dubbed the "monistic" model. According to such a model, the test of a theory *T* only requires that one compares it "with a class of facts (or observation statements) which are assumed to be 'given' somehow" (1963/1999a: 91). The monistic model, Feyerabend argued, is fatally flawed in that it neglects that there are facts "which cannot be unearthed except with the help of alternatives to the theory to be tested, and which become unavailable as soon as such alternatives are excluded" (1963/1999a: 92). In order to severely test any theory *T*, Feyerabend claimed, one should deploy *at least* one alternative to *T*. More specifically, Feyerabend put forward a pluralistic model of theory testing according to which one should operate with "*a whole set of partly overlapping, factually adequate, but mutually inconsistent theories*"

(1963/1999a: 92, italics in the original).[5] The pluralistic model of theory testing is what we refer to here as the "limited pluralism" view.

Since in 1997 the DBM certainly counted as an alternative to current approaches to the treatment of tumors, it may be tempting to think that, based on the limited pluralism view, Feyerabendian pluralists would of course have wanted to give the DBM a fair hearing, perhaps even recommending the authorization of the clinical trials. However, in 1997 the DBM lacked – just as it does today – a feature required to qualify as the kind of alternative with which one is supposed to traffic to operate Feyerabend's pluralistic model of theory testing: the DBM was not factually adequate in any reasonable sense of the notion of "factual adequacy". More specifically, since there was no publicly available evidence concerning the DBM's safety and efficacy, there was no reason to view the DBM as an alternative possessing, as per the pluralistic model, the potential to discover relevant facts – concerning either carcinogenesis or the safest and most effective way to treat tumors – that had remained hidden from view due to the dominance of chemotherapy. As Feyerabend succinctly put the point, "[n]ot all alternatives are equally suited for the purpose of criticism" (1965/1981: 109). In order for a Feyerabendian pluralist to recommend conduct of the trials of the DBM, Di Bella and his supporters would have needed (at least preliminary) evidence in hand, suggesting that the set of theories to be used in a severe test of the safety and efficacy of current tumor treatments should include the DBM. In other words, Feyerabendian pluralists would have scolded Di Bella for only partially applying Feyerabend's principle of proliferation, which recommends that one should "[*i*]*nvent and elaborate theories which are inconsistent with the accepted point of view*" (1965/1981: 105, italics in the original). In fact, while Di Bella certainly complied with the imperative to invent an alternative to the dominant theory, the DBM fared to say the least very poorly in terms of so-called "elaboration" – that is, in terms of the set of tasks required of anyone championing a self-proclaimed radical alternative to current scientific "orthodoxy", including the collection of evidence supporting the alternative. In short, then, in 1997 Feyerabendian pluralists would have voted against the authorization of the trials of the DBM.

4. *The oceanic view vis-a-vis the Di Bella case: do authorize the trials!*

Or maybe not. As is well-known, Feyerabend's understanding of theoretical pluralism became more and more permissive as time went by. For instance, in

[5] On the much discussed pluralistic model put forward by Feyerabend, see among others Preston (1997: Ch. 7), Farrell (2003: Ch. 5), Oberheim (2006), Bschir (2015), Tambolo (2015) and Collodel (2016).

the first edition of *Against Method* he explicitly suggested that one could do away with the condition of factual adequacy of the alternatives to the dominant point of view (1975: 41, note 8). And in any case, as early as 1965 he declared that the principle of proliferation "prevents the elimination of older theories which have been refuted", since such theories "contribute to the content of their victorious rivals" (1965/1981: 107). More specifically, Feyerabend championed the view that knowledge "is not a series of self-consistent theories that converges towards an ideal view; it is not a gradual approach to the truth. It is rather an ever increasing *ocean of mutually incompatible alternatives*" (1993: 21, italics in the original). On such "an extremely cumulativist account of science" (Niiniluoto 1999: 294), no alternative ever drops out of the picture:

> All theories, even those which for the time being have receded into the background, may be said to possess a "utopian" component in the sense that they provide lasting, and steadily improving, measuring sticks of adequacy for the idea which happens to be at the centre of attention (Feyerabend 1965/1981: 107).

Sure, in 1997 the DBM was not one of those old, superseded theories that, on Feyerabend's account, one should nevertheless treat as an integral part of current science: it was, rather, a newcomer. Still, nothing in theoretical pluralism as Feyerabend came to characterize it suggests that there should be barriers to entry in the market of ideas: "experts and laymen, professionals and dilettanti, truth-freaks and liars – they are all invited to participate in the contest and to make their contribution to the enrichment of our culture" (1993: 21). And Feyerabend's support not only of theoretical, but also of methodological pluralism can easily be construed as licensing precisely the harsh criticisms against current methodological conventions voiced by Di Bella. In short, then, it is easy to imagine a Feyerabendian pluralist who, in 1997, either recommends the authorization of the trials, or even suggests that in fact, no trial is needed for any drug or treatment whatsoever, since when it comes to assessing them, the attending physician's clinical eye is all that matters.

That a Feyerabendian pluralist could plausibly have supported the authorization of the trials is, in our view, a sign of the problems faced by Feyerabendian pluralism, as it evolved into a doctrine that did not include some appropriately demanding condition of adequacy for the alternatives welcome to enter the competition, such that knowledge gets characterized as an ever increasing ocean of mutually incompatible alternatives. Abandoning the very idea of factual adequacy is to say the least a very incautious move, especially in light of another relevant aspect of the case analyzed here.

In the course of 1998, after quite some resistance on the part of Di Bella had been overcome, a team of researchers operating on behalf of the Ministry

of Health, led by epidemiologist Eva Buiatti, was granted unrestricted access to Di Bella's archive and conducted a study of the records of patients therein contained – the oldest one dating back to 1971 (Buiatti *et al.* 1999). The study showed that Di Bella's claims concerning the thousands of patients that had successfully been treated with the DBM were utterly unsubstantiated. First of all, the archive included more or less three thousand items, while Di Bella had repeatedly boasted about the much higher figure of ten thousand patients. Secondly, about half of the records had to be excluded from the study, since they either did not contain a cancer diagnosis, or contained no diagnosis at all. Third, the analysis of the records of patients showed that the drugs prescribed by Di Bella to his cancer patients "changed widely" (Buiatti *et al.* 1999: 2145) during the 1971-1997 period, and that the "protocol proposed by Dr. Di Bella" for the trials "does not correspond to his prescriptions in the period 1971-1997" (Buiatti *et al.* 1999: 2148). Fourth, the survival probability observed in the patients treated with the DBM was lower than that observed in the control group, built by appropriately matching each one of the records analyzed with up to four randomly selected cases, drawn from public registries. The comparison between the observed survival among Di Bella's patients and the observed survival among cases used as a control concerned 176 patients. Of these 176 patients, only four had been treated with the DBM as the first or only therapeutic choice, and only one was alive at the end of follow-up (Buiatti *et al.* 1999: 2144-2145).

In short, then, before the 1998 phase II trials, not only was the alleged evidence of the safety and efficacy of the DBM not publicly available: there was no evidence at all – not in Di Bella's archive, nor anywhere else. Authorizing experiments on human subjects in the absence of any reason to presume the possible efficacy of the DBM was to say the least a very questionable decision, from the ethical, scientific, and economic point of view. The culprit, of course, was not Feyerabendian pluralism, but rather, the Italian Parliament. Yet, it is doubtful that Feyerabendian pluralists could have had anything to say against the authorization, unless they appealed to some condition of factual adequacy of the alternative to the dominant theory – that is, unless they stuck to the limited pluralism view and abandoned the oceanic view. Or perhaps, as we shall see in Section 5, they could have said something against authorizing the trials.

5. *Respectable thinkers* vs. *cranks*

Feyerabendian pluralists could certainly complain that our discussion above relies on the benefit of hindsight. In fact, the study of the records of patients contained in Di Bella's archive was conducted *after* the Parliament had authorized the clinical trials of the DBM, and its results were published when the

trials were well over. And in any case, Feyerabendian pluralists could retort, our discussion is uncharitable.

In fact, as many of Di Bella's critics immediately pointed out in 1997, it was easy to ascertain that despite Di Bella's grandiose claims concerning his status as a possible Nobel Prize winner, he could hardly claim expertise in the relevant scientific field: in the overwhelming majority of cases, his publications consisted of one-page abstracts of the contributed talks that he had given over the years at various conferences and workshops. Moreover, the DBM was created and perfected in the private laboratory of physiology founded and funded by Di Bella, who for decades worked in substantial isolation from the relevant scientific community: as Di Bella openly acknowledged, the only attempt to publish the results of his work on cancer treatment in a respectable peer-reviewed venue was met with a rejection letter (Di Bella *et al.* 1998: 114-115). Furthermore, Di Bella's reluctance to hand in the records of patients which, he alleged, would support his claims concerning the safety and efficacy of the DBM raised many an eyebrow, and not only among the most vocal of his critics. And finally, Di Bella's criticism of "orthodox" or "official" medicine went well beyond what one would expect to see in a properly scientific discussion of the respective merits of chemotherapy and the DBM. To mention but one instance, Di Bella's son Adolfo summarized his father's overall attitude towards the pharmaceutical industry as follows: the "'Invisibles' that shape the world's destiny" (Di Bella 2012: 307) simply cannot afford the emergence of a safe and effective alternative to chemotherapy. And to confirm that this is a pretty accurate summary of Luigi Di Bella's views, one only needs to read what he wrote in a book aimed at the lay public which came out in 1997, just a few months prior to his rise to national fame: "In this muddy and rotten environment [that of approved cancer treatments], there is the victim who suffers and dies, and there is the vampire who gets fat and prospers" (Di Bella 1997: 65, underlined in the original).

Feyerabendian pluralists can then rebut the conclusion of our discussion in Section 4 by saying that all that they would have needed in 1997-1998 to properly deal with the Di Bella case was the distinction, drawn by Feyerabend himself, "between 'respectable' people and cranks" (1964/1981: 199). In Feyerabend's words, such a distinction

> lies in the research that is done once a certain point of view is adopted. The crank usually is content with defending the point of view in its original, underdeveloped, metaphysical form, and he is not at all prepared to test its usefulness in all those cases which seem to favor the opponent, or even to admit that there exists a problem. It is this further investigation, the details of it, the knowledge of the difficulties, of the

general state of knowledge, the recognition of objections, which distinguishes the "respectable thinker" from the crank (1964/1981: 199).

Based on the information available to the public *before* the Parliament authorized the trials, Di Bella seemed to check all the boxes for a crank.[6] And based on the respectable thinker vs. crank distinction, for a Feyerabendian pluralist it would have been obvious that the right thing to do was to side against the authorization of the trials. However, there is nothing in the distinction – with which we cannot find fault – that makes it a distinctively Feyerabendian one. In other words, Feyerabendian pluralists can very plausibly claim that when it comes to spotting crankiness, they are no less well positioned than anyone else. However, recognizing a crank when they see one is hardly an ability that Feyerabendian pluralists are endowed with *qua* Feyerabendian pluralists. And there is more.

In the passage quoted above, Feyerabend draws the attention to "the research that is done once a certain point of view is adopted" as a way to tell a respectable thinker from a crank. The respectable thinker, Feyerabend suggests, is well aware of the state of background knowledge, and willing to engage in the laborious tasks required to defend and elaborate their point of view in the face of difficulties. Now of course, doing the high-quality work expected of a respectable thinker does not necessarily lead to success. For instance, serious scientific research does not necessarily lead to alternatives to the dominant point of view enjoying a high degree of success – of factual adequacy, to use the term employed by Feyerabend when describing the pluralistic model of theory testing. Still, theories exhibiting an appropriate degree of factual adequacy are precisely what one expects the work of the respectable scientist to yield, when said work is successful. A respectable scientist is supposed to champion alternatives arguably showing at least some intersubjectively appreciable merit, possibly with the potential to become, although perhaps not immediately, real contenders.

Feyerabend's emphasis on the respectable thinker vs. crank distinction at the very minimum points in the direction of the idea of alternatives possessing an appropriate degree of factual adequacy – that is, of the limited pluralism view. The view of knowledge as an ever increasing ocean of mutually incompatible alternatives espoused by Feyerabend pulls in the opposite direction. And in spite of how eloquently the oceanic view of knowledge is formulated, it is pretty impossible to square it with the fact that (not only) in the Di Bella case, the situation can be briefly described as follows: one alternative – in this

[6] Unlike other commentators of the episode, we do not view Di Bella as a fraudster, since to the best of our knowledge, he was in good-faith convinced of the actual truth of his claims concerning the efficacy of the DBM.

specific instance, the dominant approach – has been shown to be effective (unfortunately, only to a degree); the other alternative – the newly-introduced DBM – was never a candidate worth of serious scientific consideration. Therefore, we suggest, a Feyerabendian pluralist could have dealt with the Di Bella case in the correct way by adopting a properly limited or restricted form of pluralism, which requires one to drop the oceanic view of knowledge whereby no alternative ever drops out of the picture. At least since the mid-1960s, Feyerabend consistently championed the oceanic view, glorified in all editions of *Against Method* (e.g., 1975; 1993), as wedded to his pluralistic model of theory testing. Feyerabendian pluralists, we suggest, should embrace the latter but not the former.

6. *Concluding remarks*

When the Italian Parliament authorized the clinical trials of the DBM, there was no shortage of willing experimental subjects. Quite on the contrary, many who would have wanted to enroll in the trials were unable to secure a spot: when Di Bella was at the peak of his fame, many believed – or at least hoped – in the DBM's efficacy. As mentioned, an important role in the process that led to the authorization was played by the fact that an unknown but sizable number of patients elected to abandon approved treatments in favor of the DBM. Still, we maintain, the decision made by the Italian authorities was blameworthy on the ethical, scientific, and economic level. Our discussion shows that a Feyerabendian pluralist would plausibly have found themselves in agreement with such a decision, unless they dropped the oceanic view of knowledge and stuck to a limited version of pluralism revolving around the pluralist model of theory testing.

The gist of the present analysis, it seems to us, is in line with some important threads emerged in recent Feyerabend scholarship. In fact, it is well-known that at least in some of his later writings, Feyerabend openly acknowledged the resistance that the world offers to some of our attempts to describe it. As he put it, some approaches to the world "find no point of attack in it and simply collapse" (1999b: 145). Unfortunately, Feyerabend was never too explicit on how one should proceed to strike the right balance between the acknowledgment of said resistance and his pluralist leanings (see, e.g., Tambolo 2014; Brown 2016; Giere 2016). However, Feyerabendian pluralists can do better than Feyerabend himself did in this regard. This is attested, to mention but one instance, by recent approaches to the issue of science-based policy advice concerning such urgent matters as the COVID-19 pandemics that explicitly qualify themselves as "inspired" (Bschir *et al.* 2022) by Feyerabend's plural-

ism – that is, we maintain, by the limited, restricted pluralism that would have allowed a Feyerabendian pluralist to properly deal with the Di Bella case.

Luca Tambolo
Independent scholar
ltambolo@gmail.com

Gustavo Cevolani
IMT School for Advanced Studies Lucca
gustavo.cevolani@imtlucca.it

Acknowledgments: We are grateful to the participants and the audience of the worskhop "The Role of Experts in Democratic Societies. In Honor of Paul K. Feyerabend" (IMT School for Advanced Studies Lucca, June 16, 2023) for questions and suggestions and to Roberto Festa for his comments on a draft of the paper. Gustavo Cevolani acknowledges financial support from the Italian Ministry of University and Research (MUR) through the PRIN 2022 grant n. 2022ARRY9N (Reasoning with hypotheses: Integrating logical, probabilistic, and experimental perspectives), funded by the European Union (Next Generation EU).

References

Brown, Matthew J., 2016, "The Abundant World: Feyerabend's Metaphysics of Science", in *Studies in History and Philosophy of Science*, 57: 142-154.

—, 2021, "Against Expertise. A Lesson from Feyerabend's *Science in a Free Society*?", in Bschir, Karim, Jamie Shaw, eds., *Interpreting Feyerabend. Critical Essays*, Cambridge University Press, Cambridge: 191-212.

Bschir, Karim, 2015, "Feyerabend and Popper on Theory Proliferation and Anomaly Import: On the Compatibility of Theoretical Pluralism and Critical Rationalism", in *HOPOS*, 5: 24-55.

—, Simon Lohse, 2022, "Pandemics, Policy, and Pluralism: A Feyerabend-Inspired Perspective on COVID-19", in *Synthese*, 200: 441.

Buiatti, Eva *et al.*, 1999, "Results from a Historical Survey of the Survival of Cancer Patients Given the Di Bella Mulitherapy", in *Cancer*, 86: 2143-2149.

Collodel, Matteo, 2016, "Was Feyerabend a Popperian? Methodological Issues in the History of the Philosophy of Science", in *Studies in History and Philosophy of Science*, 57: 27-56.

Di Bella, Adolfo, 2012, *Il poeta della scienza. Vita del professore Luigi Di Bella*, Emmeci Edizioni, Rome.

Di Bella, Giuseppe, Adolfo Di Bella, Luciano Gualano, 2012, "Laudatio for the Centenary of the Birth of Luigi Di Bella, MD, PhD", in *Neuroendocrinology Letters*, 33: 247-248.

Di Bella, Giuseppe, 2010, "The Di Bella Method (DBM)", in *Neuroendocrinology Letters*, 31: 7-42.

Di Bella, Luigi, 1997, *Cancro: siamo sulla strada giusta?*, Travel Factory, Rome.

—, Bruno Vespa, 1998, *Si può guarire? La mia vita. Il mio metodo. La mia verità*, Mondadori, Milan.

Ernst, Edzard, 2021, *So-Called Alternative Medicine (SCAM) for Cancer*, Springer, Cham.

Farrell, Robert, 2003, *Feyerabend and Scientific Values. Tightrope-Walking Rationality*, Kluwer, Dordrecht.

Feyerabend, Paul Karl, 1963/1999a, "How to Be a Good Empiricist. A Plea for Tolerance in Matters Epistemological", as reprinted in Feyerabend 1999a, pp. 78-103.

—, 1964/1981, "Realism and Instrumentalism: Comments on the Logic of Factual Support", as reprinted in Feyerabend 1981: 176-202.

—, 1965/1981, "Reply to Criticism. Comments on Smart, Sellars and Putnam", as reprinted in Feyerabend 1981, pp. 104-131.

—, 1975, *Against Method. Outline of an Anarchistic Theory of Knowledge*, New Left Books, London.

—, 1978, *Science in a Free Society*, New Left Books, London.

—, 1981, *Realism, Rationalism and Scientific Method. Philosophical Papers, Volume 1*, Cambridge University Press, Cambridge.

—, 1987, *Farewell to Reason*, Verso, London.

—, 1993, *Against Method*, 3rd ed., Verso, London.

—, 1999a, *Knowledge, Science and Relativism. Philosophical Papers, Volume 3*, edited by John Preston, Cambridge University Press, Cambridge.

—, 1999b, *Conquest of Abundance. A Tale of Abstraction Versus the Richness of Being*, edited by Bert Terpstra, The University of Chicago Press: Chicago.

Giere, Ronald, 2016, "Feyerabend's Perspectivism", in *Studies in History and Philosophy of Science*, 57: 137-141.

Italian Study Group for the Di Bella Multitherapy Trials, 1999, "Evaluation of an Unconventional Cancer Treatment (the Di Bella Multitherapy): Results of Phase II Trials in Italy", in *British Medical Journal*, 318: 224-228.

Kidd, Ian James, 2016a, "Why Did Feyerabend Defend Astrology? Integrity, Virtue, and the Authority of Science", in *Social Epistemology*, 30: 464-482.

—, 2016b, "How Should Feyerabend Have Defended Astrology?", in *Social Epistemology Reply and Review Collective*, 5: 11-17.

Lloyd, Elizabeth, 1997/2000, "Feyerabend, Mill, and Pluralism", as reprinted in Preston, John, Gonzalo Munévar, David Lamb, eds., 2000, *The Worst Enemy of Science?*, Oxford University Press, Oxford: 115-124.

Markle, Gerald E., James C. Petersen, eds., 1980, *Politics, Science, and Cancer: The Laetrile Controversy*, Westview Press, Boulder.

Minutes of the January 14, 1998, Parliamentary Hearing, Committee for Social Affairs of the Italian Chamber of Deputies, accessed February, 2024, http://leg13.camera.it/chiosco.asp?cp=2&content=/_dati/leg13/lavori/stencomm/tabindag.htm.

Minutes of the January 15, 1998, Parliamentary Hearing, Committee for Social Affairs of the Italian Chamber of Deputies, accessed February, 2024, http://leg13.camera.it/chiosco.asp?cp=2&content=/_dati/leg13/lavori/stencomm/tabindag.htm.

Minutes of the January 20, 1998, Parliamentary Hearing, Committee for Social Affairs of the Italian Chamber of Deputies, accessed February, 2024, http://leg13.camera.it/chiosco.asp?cp=2&content=/_dati/leg13/lavori/stencomm/tabindag.htm.

Niiniluoto, Ilkka, 1999, *Critical Scientific Realism*, Oxford University Press, Oxford.

Oberheim, Eric, 2006, *Feyerabend's Philosophy*, de Gruyter, Berlin.

Pigliucci, Massimo, 2016a, "Was Feyerabend Right in Defending Astrology? A Commentary on Kidd", in *Social Epistemology Reply and Review Collective*, 5: 1-6.

—, 2016b, "How Should Feyerabend Have Defended Astrology? A Further Reply to Kidd", in *Social Epistemology Reply and Review Collective*, 5: 10-16.

—, 2016c, "Feyerabend and the Cranks: A Reply to Shaw", in *Social Epistemology Review and Reply Collective*, 6: 1-6.

Preston, John, 1997, *Feyerabend: Philosophy, Science and Society*, Polity Press, Cambridge.

Shaw, Jamie, 2017, "Feyerabend and the Cranks: On Demarcation, Epistemic Virtues, and Astrology", in *Social Epistemology Review and Reply Collective*, 6: 74-88.

—, 2021, "Feyerabend and Manufactured Disagreement: Reflections on Expertise, Consensus, and Science Policy", in *Synthese*, 198: 6053-6084.

Tambolo, Luca, 2014, "Pliability and Resistance: Feyerabendian Insights into Sophisticated Realism", in *European Journal for Philosophy of Science*, 4: 197-213.

—, 2015, "A Tale of Three Theories. Feyerabend and Popper on Progress and the Aim of Science", in *Studies in History and Philosophy of Science*, 51: 33-41.

Expanding epistemic public trust. What role for expert-lay communication?

Piero Avitabile, Alessandro Demichelis

Abstract: This paper examines how communication between experts and lay citizens influences the development of criteria for epistemic public trust, building upon the detailed framework proposed by Irzik and Kurtulmus (2019). We first analyse the epistemic significance of trust and its implications in the public sphere. Our focus is twofold: identifying what attributes make experts trustworthy and exploring the reasons and second-order evidence that lay people can utilize to justify their trust in experts. We argue that the way experts engage in argumentation plays a crucial role: it is essential for their epistemic responsibility towards lay citizens and serves as a key indicator of their trustworthiness. Based on these considerations, we suggest an additional criterion for experts – facilitating rather than hindering the public discussion. Using an example from expert and layperson discussions during the COVID-19 pandemic, we demonstrate how this criterion can be practically applied.

Keywords: Feyerabend, Trust, Epistemic trust, Scientific communication, Argumentation theory.

1. *Introduction*

The advice in all cases is to use experts, but never to trust them and certainly never to rely on them entirely (Feyerabend 1978: 97).

What did Feyerabend intend to say with this footnote included in *Science in a Free Society*? He certainly had in mind a difference between use, trust, and reliance. Why, though, does he advise not to trust experts? Is there anything that experts can do in order to demonstrate their trustworthiness to someone who is inclined to listen to the advice of the Austrian philosopher?

In our article, we intend to elaborate a proposal in line with the spirit, if not the letter, of Feyerabend's ideas. We aim to analyse the relationship between experts and laypeople when the flow of information is involved, expanding on Irzik and Kurtulmus' recent proposal (2019) of a set of criteria to be followed to foster epistemic public trust in science.

We propose a new criterion focused on the need for experts to facilitate

philinq XII, 1-2024, pp. 163-179
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v12i1.557
Submitted: 4 June 2024
Accepted: 22 July 2024

rather than hinder further inquiry from laypersons. To justify this addition, we proceed as follows.

In Section 2, we draw on existing social epistemology literature to clarify the concept of trust, defining its boundaries and characteristics as an epistemological notion. In Section 3, we present Irzik and Kurtulmus' (2019) criteria for warranted epistemic public trust. In Section 4, we introduce various "tools" that have been suggested for laypeople to assess conflicting expert testimonies. In Section 5, we present the core of our proposal, suggesting that Irzik and Kurtulmus' (2019) criteria are adequate but incomplete, in that they do not fully account for a crucial dimension of epistemic trustworthiness, namely *epistemic responsibility* (Anderson 2011). In particular, we hold that useful indicators of experts' epistemic responsibility come from their argumentative conduct in public communication. Therefore, we propose that another criterion should be added, one that takes into account the necessity, when public interests are at stake, that an expert testimony does not hamper the public discussion of the relevant topics in the affected community of inquirers.

We cannot know what Feyerabend would make of our proposal. However, in line with him, we think that from epistemic asymmetry need not follow a hierarchy of power, and that citizens have an important role in treading the path that leads to the creation of knowledge and the resolution of public problems. We aim to demonstrate a viable pathway.

2. *Trust and its epistemic side*

2.1 *The epistemic side of trust*

Our proposal is the addition of a new criterion necessary to the construction of public epistemic trust. But in order to do so we have first to clarify what we mean with public epistemic trust. We will briefly consider the notion of *trust*, subsequently expanding to the literature concerning *epistemic* trust, and finally considering the notion of *public* epistemic trust.

First of all, we have to canvass a brief description of the salient features of *trust.* In line with an ample philosophical tradition (Baier 1995; Zagzebski 1996, 2012) we structure trust as a triadic relation.[1] The three elements of that relation are the trustor, the trustee, and what is entrusted by the trustor to the trustee. The actual nature of this third element is as broad as possible: it could be an action, a statement, an attitude, etc. For what regards the structural fea-

[1] For an overview of the most recent developments in the discussion over trust and trustworthiness, see Lalumera (2024).

tures of trust, four main aspects of trust can be identified: reliance, vulnerability, confidence, and expectation (McCraw 2015).[2] Reliance, or dependence, is often taken as the most important aspect of trust: trusting someone means relying on that someone. Reliance leads to the possibility of being let down, an aspect where vulnerability emerges: trusting means voluntarily leaving oneself exposed to risk. This feature, in turn, paves the way to confidence, which is a "distinctive, and affectively loaded, way of seeing the one trusted" (Jones 1996).[3] Finally, confidence expresses itself in the possibility of formulating a prediction regarding the behaviour of the trustee.

These elements are all required to obtain a definition of trust, but none of them is sufficient on its own. Consider some examples. In a case, we see Joe, a gambler with a huge debt from shady sources, betting everything he owns on the underdog, hoping to repay his debt and avoid a gruesome retaliation (McCraw 2015). Here we can say that Joe relies on his bet, but we cannot say that he trusts the underdog to win, due to a lack of confidence. In another instance, the alarm fails to ring in the morning, and Jane arrives late at work. Jane relied on the alarm to ring, but did not exert emotional affection on the clock. Therefore, she might be upset by the malfunctioning, but she does not feel her trust betrayed.

From trust in general, we shall now narrow our gaze to a more specific kind of trust: trust in someone *as a provider of information*. This type of trust is often called *epistemic trust* (Wilholt 2013). Epistemologists have not always given the dignity of consideration to the concept of trust. Nowadays, however, trust is increasingly granted attention as an epistemic notion (Origgi 2004). In the contemporary world, the cumulative and ever-expanding body of knowledge and the multiplication of actors capable of its production and dissemination (universities, public and private-funded research groups, governments, media) means that it is impossible for a given individual to obtain first-hand, direct knowledge for every piece of information needed. Thus, trust has been increasingly considered as playing a fundamental role in the construction of knowledge. As John Hardwig notices in a seminal paper, the epistemic stance that "trusting and knowing are deeply antithetical" cannot be kept in a world

[2] It is important to notice that for McCraw (2015) the two main features of trust are reliance and confidence, whereas vulnerability and expectations are secondary features that emerge from, respectively, reliance and confidence. An extended discussion about the relative weight of each component, however, falls outside the scope of this work.

[3] Following the literature, we use the term "confidence" to identify a specific feature of trust. However, it has to be noticed that the word "confidence" comes from the Latin *confidentia*, which means "trust", a relationship that remained in the modern French cognate *confiance*, which again means "trust". Therefore, it might as well be just a truism that "trust involves confidence" (Smolkin 2008).

where greater specialization requires scientific perspectives to be narrower in order to be more acute, and therefore, "those who do not trust cannot know" (1991: 639). According to Hardwig, trust is justified when there is a particular class of reasons that grounds the acquisition of knowledge through testimony. These reasons refer to the assessment of moral and epistemic qualities of the testimony, which in this case is the trustee. While the assessment of the moral side is already addressed in the general account of trust, competence (related to a specific task or domain) is the distinctive feature of epistemic trust. "Epistemic trust is trust, on the one hand, in the goodwill of others, and, on the other hand, in their competence" (Origgi 2004: 64).

How the presence of the element of competence distinguishes epistemic trust from general trust can be elucidated with an example. Consider how I trust my mother implicitly and I am sure of her goodwill toward me; at the same time, I would not ask her to change the oil of my motorcycle, a task that relies on being sufficiently competent not to incur costly mistakes. Conversely, the relevance of the competence requirement does not mean that the goodwill element becomes secondary. I would not ask to perform an oil change to a mechanic I know to be competent, but also fishy and prone to scam customers. Epistemic trust is given to an agent that we consider capable and competent, and that we also judge to act on that competence with benevolence.

Epistemic trust plays a particularly important role in science, and its role in scientific research within and across scientific communities has been of central interest in social epistemology. Its mediating role becomes evident when considering how the division of epistemic labour (Kitcher 1990; Brennan 2010) underpins the specialization and the disciplinary divisions of the scientific community. No chemist would obtain much without the knowledge produced by physics, no neuroscientist would be able to work without the foundation laid down by physiology and electromechanics. But at the same time, no scientist can have specialized expertise in those disciplines that fall outside their focus and background preparation. In that context, other non-epistemic factors come into play when grounding epistemic trust.

2.2. *Inductive risk and epistemic trust*

In particular, the level of trade-off between the benefit of correct results versus the cost of incorrect ones is a value judgment, one that differs from discipline to discipline and even from scientist to scientist (Douglas 2000; 2009). In philosophy of science, this trade-off is often referred to as the question of inductive risk. To paraphrase Rudner (1953), no scientific hypothesis is ever permanently verified; therefore, in deciding to accept a hypothesis a scientist *qua*

scientist must make value judgments. The reason is that accepting a hypothesis requires at the same time having established a threshold for its acceptance. This level of acceptance changes depending on the gravity of the consequences that result from accepting a false hypothesis or rejecting a true one. For example, imagine a quality control manager who has to test the hypothesis that their products are safe. The evidential standards for testing their hypothesis will sensibly change whether the products in question are life-saving medicines or buckles. The result of this weighting, however, depends on non-epistemic values (such as the importance one attaches to saving lives) that are therefore inextricably tied to the very production of scientific knowledge (Elliott *et al.* 2014).

Understanding that non-epistemic values such as moral, personal, social, political and cultural values can influence and shape scientific research is intuitively easy. However, this influence can weaken the universality and applicability of scientific results. Therefore, a normative requirement has been proposed: science should strive to minimize or eliminate the non-epistemic values from their research (Reiss *et al.* 2020). This proposal has been criticized as unattainable, due to the unavoidable presence of non-epistemic values in determining the level of error in acceptance or rejection of hypotheses. If the inductive risk argument is valid (and we believe it is), then assigning trust to scientific experts also requires lay people to trust the value judgments that are inherent to experts *qua* experts. Some authors remark (Wilholt 2013) that sharing the same value judgments in the set of utilities used to weigh the outcome of research is what constitutes the difference between epistemic reliance and epistemic trust, at least for what trust between scientists is concerned. This element allows us to underline the fact that the topic of epistemic trust has been treated, in the analysis carried out within the context of social epistemology, as something that is chiefly needed to ground the knowledge *within* the boundaries of the scientific discourse, i.e. in the dialogical exchange between experts of different disciplines. An extension of the discussion that would also involve the experts-laypeople relationship started as a topic belonging to different research programs, such as the "Strong Program" of sociology of science (Latour 1986), or the initiatives aimed at increasing the level of Public Understanding of Science (Short 2013). However, the epistemic relationship between experts and non-experts has increasingly received interest from epistemologists as well. Alvin Goldman's (2001) stands out as a normative analysis of what he calls the *novice/2-experts* problem, where a person who is not an expert confronts a situation where two different putative experts diverge in their claims. To ground a form of discernment, Goldman proposes a set of criteria for identifying the most trustworthy expert, even without having specific expertise (we elaborate on these criteria in Section 4).

Communication as a socio-epistemic practice seems to deserve particular attention when it comes to its role in shaping trust relations between different epistemic communities. Quite surprisingly, however, communication between experts and laymen has not often been explicitly thematized for its role in epistemic trust. In particular, we would like to follow McCraw (2015) in the idea that "there must be something about [an expert] that communicates the content of [a] belief [...] to [a layperson] so that [that layperson] can have the belief in question. Thus, there must be some kind of communication component here to explain or ground how the truster comes by his/her belief" (McCraw 2015: 8). Communication is indeed a tricky business, but one full of consequences for the assignment of epistemic trust. Trusting someone's word leaves one open to vulnerability, in terms of "deception, incompetence, or false assertions" (2015: 10), as well as miscommunication, ambiguity, and simple misunderstanding. Expanding the discussion over what features should show effective lay-expert communication is central to the aim of the next sections, where we provide an account of epistemic public trust.

3. *Criteria for epistemic public trust*

In the light of what has been said so far on epistemic trust (ET), we are ready to discuss more systematic attempts at normatively capturing it in its public dimension. We take the public dimension of epistemic trust to be accounted for by the epistemic asymmetry that occurs between the trustor and the trustee, and the nature of the object of trust. In its most simple form, the trustor can be understood as a layman, the trustee as an expert, and the object of trust as an expert testimony of public interest. However, this representation will often be too simplistic. In a stronger and more relevant sense of the qualification "public", the trustor can also be an expert while the trustee is a layman, as it is clear in cases where citizens possess some form of specialized relevant knowledge.[4] Public trust is also better described as a social rather than an individual matter, where groups of citizens trust groups of experts (as the very title of Contessa 2023 says: *it takes a village to trust science*), and vice versa. Furthermore, the public is not to be considered as a homogeneous entity (cf. Resnik 2011; Irzik *et al.* 2019). Rather, depending on the case at hand and the interests at stake, there will be a relevant community identified as the target audience.

For the sake of simplicity, but with these considerations in mind, we will focus on the more individual dimension of epistemic public trust (EPT), devoting most of our discussion to the layman-to-expert direction of this relation.

[4] On the matter of the bidirectionality of trust, see Barrotta and Montuschi (2019) and footnote 8.

In doing so, we are also confident that the normative demands that will emerge will be generalizable to the social dimension, despite not fully accounting for it.

For such a framework to be normative, requirements need to be set both for experts' trustworthiness and laypeople's warranted assignment of epistemic trust. A systematic attempt at providing such a normative account of EPT can be found in Irzik and Kurtulmus (2019). In the next pages, our considerations and remarks are going to build and expand on their proposal, which we will use as a starting point for our discussion. At the core of their account, there are six criteria for what they term "enhanced" warranted epistemic public trust, which are the following:[5]

(1) E believes that *p* and honestly communicates it to L.
(2) L takes E's testimony *p* as a (strong but defeasible) reason to believe that *p*.
(3) *p* is the output of reliable scientific research carried out by E.[6]
(4) L relies on E because she has good reasons to believe that (3).
(5) When public welfare is at stake, in making methodological decisions regarding the distribution of inductive risks with respect to *p*, E makes those decisions in agreement with L's assessments of the inductive risks in question.
(6) L has reason to believe that (5) is satisfied.

According to Irzik and Kurtulmus, then, "the relevant public will be said to invest warranted epistemic trust in [E] as a provider of [*p*] when a sufficient number of its members invest warranted epistemic trust in [E] as a provider of [*p*]" (2019: 1150).

The structure of these criteria is symmetrical. On the one hand, criteria (1), (3), and (5) provide requirements for expert trustworthiness. In particular, (1) and (3) respectively account for honest communication and reliability of the testimony. For this account to be generally valid, more fundamental assumptions need to already be in the background. For example, the competence component is only accounted for by (3), which takes into account the reliability of the process that gives *p* as an output. However, nothing is said about *E* being epistemically well-placed with respect to *p*, e.g. by being a genuine expert in the field *D*.

[5] The criteria listed have been slightly modified for consistency and clarity. "Enhanced" is opposed to "basic" epistemic public trust, which only comprises criteria (1) to (4).

[6] As rightly pointed out by one of the reviewers, this criterion appears to be too restrictive, as it does not allow an expert to rely on other experts' knowledge. We will commit to a more charitable interpretation of this criterion, and understand it as follows: "*p* is the output of reliable scientific research carried out by E *or by the community of research to which E belongs*".

To explain criterion (5), Irzik and Kurtulmus (2019) recur to the following example. Imagine a teacher submitting a test to her students. They are split into two rooms, and in each one there is a teaching assistant supervising. At the end of the exam, both TAs come to the teacher to report that one of the students in their room was caught cheating. Now, the teacher is aware that the first TA is very cautious, and would hardly ever falsely accuse someone of cheating if they were unsure. On the other hand, the second TA is known for being very strict about not wanting anybody to get away with cheating, even if it happens at the cost of falsely accusing someone who was just behaving ambiguously. As a teacher, you will be more inclined to assign epistemic trust to one or the other TA, according to your values and the nature of the risks you are willing to take. In the analogy, the TAs are the experts while the lay citizen is the teacher. At the core of this criterion there is then the idea that experts and lay citizens need to come to an agreement on the risks at stake in falsely taking the content of a testimony to be true.

On the other hand, (2), (4), and (6) are criteria for the layman's assignment of trust. In all of these cases, the assignment is *warranted* if *L* has "strong but defeasible reasons" to deem the expert trustworthy in the relevant aspects that qualify her as such. What is deemed as a normative requirement for experts' trustworthiness (criteria 1, 3, 5) shall symmetrically have a counterpart that accounts for lay reasons for trusting experts. Irzik and Kurtulmus (2019) do not explore what those reasons are, but research in social epistemology displays a full range of them, in the form of various indicators and sociological proxies. The nature of these reasons is a matter of more detailed discussion in the next sections.

4. *Second-order evidence for trusting experts*

Since the lay public cannot, by definition, judge expert testimonies firsthand, they will have to resort to external indicators and sociological proxies. Reliability indicators tell something either about the testimony or the source of that testimony. Indicators at the testimony-level are primarily related to the consensus of the relevant scientific community (Anderson 2011); conversely, the existence of expert disagreement counts as a higher-order level of evidence working against *p* (Sosa 2011). However, there are some practical issues that can undermine the success of consensus as an indicator of the reliability of *p*. The main one is that, except for historically consolidated scientific evidence (which is rather uncontroversial), dissent is a physiological and healthy component of scientific discourse. Luckily, useful proxies of source reliability are generally easier to retrieve and assess. So, although the public is generally expected to accept (or reject) the particular testimony *p* as part of their beliefs,

their evaluation will usually focus on the reliability of the sources providing it (the primary source being the expert), and will then generalize it to the testimony *per se*.

Therefore, a hierarchy of characterizing features and proxies of expertise comes into play when laymen are called to evaluate testimonies. Consistently with the dimensions of epistemic trust just seen, the competence and honesty of experts are the two macro-categories found in the philosophical and sociological literature discussing expertise for source-level assessments (e.g., Goldman 2001; Croce 2019; Martini 2019; Watson 2020; Collins *et al.* 2019; Walton 2010). Another dimension, epistemic responsibility, is introduced by Anderson (2011), and will be more extensively discussed later on.

Whether a scientist *qua* scientist is competent and honest will depend on how reliable the social mechanisms of science in general, and/or of the particular scientific field of interest, are in selecting competent and honest scientists: a sovra-ordinate dimension of trust in these social mechanisms is therefore critical for source-level reliability assessments (Contessa 2023). Connected to this dimension is, for example, the reliability of scientific journals, where first-hand "intra-scientific" testimonies (Gerken 2022) are reported. Also dependent on the social mechanisms of science is the epistemic status of the disciplines to which the experts belong. Crucially, this accounts for the different reliability of two alleged epistemic authorities, both honest and equally competent in their respective fields, where one is a quantum physicist, and the other an astrologist. Only the former belongs to a discipline that has earned a respectable epistemic status within the institutional mechanisms of science. On the other hand, in the latter case, we would not want to qualify the honest and competent astrologer as an expert instead.

The honesty component of an expert can be evaluated through their disclosed conflicts of interest, any history of academic misconduct, and their personal integrity as observed in their previous public dealings. As for competence, the educational background and academic credentials of an expert, ranging from bachelor's degrees to PhDs and further academic advancements, are easily accessible to the general public.

Goldman (2001) proposes evaluating an expert's past track record when their testimony's truthfulness is verifiable by novices within a reasonable timeframe, as with certain predictions in astronomy, meteorology, or finance. However, this approach is mainly applicable to fields where expertise is predominantly about making accurate predictions. Goldman further recommends seeking the opinion of another expert or a meta-expert to assess the reliability of the initial expert. This strategy is especially relevant in cases of "epistemic trespassing" (Ballantyne 2019), where legitimate experts offer testimonies out-

side their area of specialization. However, this approach reintroduces concerns about the second expert's competence and honesty.

The entire field of *scientometrics* is devoted to developing metrics that can serve as indicators of scientists' productivity and influence (e.g., the *h-index*), or of how impactful the journals where they usually publish are (e.g., the *impact factor*). In this case, it is thought that laymen can get a good idea of how to interpret these metrics and implement them in their reliability judgments with little information. However, this is rarely the case, as these metrics are largely variable across and within disciplines, and can only be meaningfully interpreted within the relevant scientific community. On top of that, the existence of these metrics has often been criticized for privileging quantity over quality and fostering the "publish or perish" culture in academia. According to this line of argument, the current incentives for publishing are detrimental to the very process of trustworthy reliable scientific research (Fanelli 2010). Moreover, scientific journals are not the common source from which the public comes to know expert testimonies, which are usually mediated by reporting sources (media, scientific journalists, politicians, etc.). This intermediation in turn raises the issue of the reliability of second-order sources, again running into problems of competence and honesty. This necessitates the use of additional proxies and indicators, further complicating and obscuring the process.

5. *From epistemic responsibility to the role of public communication*

So far, we have explored the requirements for an expert to be considered trustworthy, focusing on honesty (criterion 1) and competence (criterion 3). Additionally, we have discussed valid reasons for laypeople to place their trust in experts, specifically through second-order evidence of the expert's competence and the reliability of the scientific process, as required by criteria (2) and (4).

However, the dimensions of competence and honesty do not fully account for the definition of EPT found in Irzik and Kurtulmus (2019). The last pair of criteria (5 and 6), which are at the core of the public dimension of ET, are clearly not entirely reducible to either the honesty or the competence dimension. This opens the way for the third dimension of expert trustworthiness that has only been briefly mentioned so far: epistemic responsibility, as identified in Anderson (2011).

Often overlooked, this dimension is indeed quite relevant when considering aspects like experts' openness to change their minds, to listen to lay people's needs and concerns, and their willingness to share relevant information accordingly. In Anderson's words, "the mark of epistemic responsibility is responsive accountability to the community of inquirers" (2011: 146). Being "ac-

countable" to the community of inquiry, broadly conceived as a community of experts and lay people working together to solve public problems, is crucial when the interests and values of the community itself are at stake.

How can this dimension emerge in the form of second-order evidence for lay people? Irzik and Kurtulmus (2019) provide answers that are only partially convincing: they propose that lay people should examine the design of decision-making forums, and prefer testimonies emerging from hybrid forums. In settings like consensus conferences, for example, the dialogue between experts and non-experts takes place by design (on consensus conferences, see Solomon 2015). A further reliable indicator can be the diversity within expert committees. Building on Douglas (2009), Irzik and Kurtulmus (2019) argue that decisions made by groups diverse in gender, ethnicity, and class are likely to be more objective. However, considering the infrequency of consensus conferences and the unclear impact of diversity in addressing particular community problems, it may be useful to explore more compelling second-order reasons.

Our claim is that a major part of what constitutes evidence for epistemically responsible experts is grounded in their argumentative conduct in the relevant settings of public communication. In the accounts of trust so far considered, communication has only appeared collaterally. In criterion (2) the expert is required to honestly *communicate* what she believes; even more importantly, the central point of criterion (5) is the *agreement* between E and L on the inductive risks in question. The process that allows this agreement to be reached is also relevant for the warranted assignment of epistemic trust and part of what makes an expert epistemically responsible, and still needs to be accounted for in the framework of Irzik and Kurtulmus (2019). An improper use of the expert's epistemic (and political) power to underestimate or even silence legitimate lay concerns, or the absence of recognition of relevant forms of lay expertise in solving the public problems at stake, will instead hamper this process. Therefore, the account provided in Irzik and Kurtulmus (2019) not only calls for a public setting where experts and lay citizens exchange their reasons: more fundamentally, it requires rules that govern this exchange of reasons in order to reach an agreement. These, we claim, will have to be part of the criteria for the "responsive accountability" that allows to preserve and enhance EPT, and whose rationale is rooted in the dimension of epistemic responsibility.

We propose to assess the argumentative conduct of an expert-lay dialogue according to the following general criterion, which is meant as an ideal prosecution of the six criteria provided in Irzik and Kurtulmus (2019):

(7) When public interests are at stake, E's argumentative moves should not hamper the public discussion in the relevant community of inquiry.

What is entailed by this broad formulation is a well-established idea in the pragma-dialectical approach to argumentation theory (van Eemeren *et al.* 2004). According to this view, fallacious arguments in a discussion are moves that hamper the resolution of the disagreement between the discussants. Applied to public settings, with this criterion we want to suggest the idea that experts ought to abide by rules of reasonable argumentation when discussing with non-experts. How to operationalize this criterion in the unusual setting of expert/layman dialogues?

Once again, Goldman (2001) and Anderson (2011) come to our aid. Linking it to the dimension of competence alone, Goldman thus addresses the indirect argumentative justification of experts as a possible criterion:

> Whenever expert E_2 offers evidence for her conclusion, expert E_1 presents an ostensible rebuttal or defeater of that evidence. On the other hand, when E_1 offers evidence for her conclusion, E_2 never manages to offer a rebuttal or defeater to E_1's evidence. Now N is not in a position to assess the truth-value of E_1's defeaters against E_2, nor to evaluate the truth-value or strength of support that E_1's (undefeated) evidence gives to E_1's conclusion. For these reasons, E_1's evidence (or arguments) are not directly justificatory for N. Nonetheless, in "formal" dialectical terms, E_1 seems to be doing better in the dispute. Furthermore, I suggest, this dialectical superiority may reasonably be taken as an indicator of E_1's having superior expertise on the question at issue. It is a (non-conclusive) indicator that E_1 has a superior fund of information in the domain, or a superior method for manipulating her information, or both (2001: 95).

So, according to Goldman, experts' argumentative support to their testimonies and compliance in meeting the burden of proof can serve as fallible proxies of their competence.

On the other hand, Anderson ties the criterion of "dialogic irrationality" to epistemic responsibility. She characterizes it as "continuing to repeat claims after they have been publicly refuted by others, without responding to the refutations" (2011: 147). She imagines an evolution denialist arguing with a paleontologist on the existence of transitional fossils. In the discussion, the evolution denialist does not respond in a meaningful way to the paleontologist, but merely repeats what they had previously stated, thus leading the discussion to a forced stalemate.

In both Goldman's and Anderson's views, argumentative conduct is a relevant criterion for choosing between two disagreeing experts. Building on these considerations, we hold that this criterion can be further elaborated to be beneficial for a wider range of cases: not only expert disagreements but, more importantly, also discussions between experts and lay citizens. Such interactions are not uncommon, as they occur in settings like television talk shows,

courtroom proceedings, medical consultations, government hearings – among others. These settings are of course characterized by an epistemic asymmetry; however, this does not mean that all the relevant information is on the experts' side. Lay people have privileged epistemic access to their values and interests, to their personal attitudes toward risk, and may have some relevant form of local knowledge and lay expertise (Wynne 1996; Barrotta *et al.* 2018).[7] Most importantly, lay people can also have access to testimonial evidence, i.e. what other experts have said. Naturally, the report of an expert testimony by a lay man will not be as strong as first-order evidence directly reported by relevant experts. Nonetheless, weak evidence should still be allowed in the "public court" if the stakes are judged high enough by the relevant public.

Let's consider an example of such interactions in relation to our criterion. The following dialogues are taken from an Italian talk show during COVID-19 pandemics, where an expert (Andrea Crisanti) and a politician (Claudio Borghi) disagree on lockdown policies:[8]

(a)

E: Lockdown works and saves lives.
L: Professor Ioannidis, the top expert on epidemiology, tells us that lockdown in all countries showed little to no benefit.
E: The nullity of what you said speaks for itself.

(b)

L: You are not a relevant expert. You are an expert on the Anopheles mosquito, but not on human viruses.
E: I am a relevant expert. I know exactly what genetic and environmental factors contribute to transmitting the disease. This is because the epidemic is independent of the pathogen. The epidemic, whether it is malaria, cholera or influenza, has ecological characteristics that have nothing to do with the agent that transmits them.

[7] Barrotta *et al.* (2018) discuss the case of the Vajont disaster, where almost 2000 people were killed by a massive landslide in northeast Italy in 1963. One of the tallest arch dams in the world was constructed there in 1959, despite the fact that locals cautioned scientists about the instability of the valley slopes. Their local knowledge and perception of risk were disregarded by scientists, who consequently made an epistemologically unsound judgment that led to disastrous consequences.

[8] Piazza Pulita, 22.04.2021. The dialogue is here reconstructed and modified for simplicity. It can be retrieved at: https://www.la7.it/piazzapulita/video/lo-scontro-a-piazzapulita-tra-borghi-e-il-prof-crisanti-preferisco-sentire-i-massimi-esperti-al-22-04-2021-376930. Interestingly, a response from Prof. Ioannidis was prompted after a week by the same TV program, and the expert corrected what Borghi attributed to him. This clearly shows the role that journalists can play in fostering EPT: https://www.la7.it/piazzapulita/video/lintervista-integrale-al-prof-ioannidis-che-risponde-a-borghi-sui-lockdown-29-04-2021-378242.

In case (a), the layman brings forward some testimonial evidence (viz., expert I's testimony that *p*). E does not address the testimony reported by the layman, and in doing so he makes a sort of appeal to the layman's ignorance (an argument along the lines of: since you are not an expert in *D*, what you are saying is not true). This move is fallacious in that it does not support the disputed standpoint with a relevant reason, forcefully halting the exchange of reasons on that point, thus hindering the continuation of the public discussion. In case (b), there is again disagreement between E and L, but E provides reasons to back his claim. Only this latter case provides the layman with second-order reasons to trust the expert. Not only does it serve as an indicator of expert competence (the expert shows that they know what they are talking about, as they can provide reasons for it), but also (and more crucially) for their epistemic responsibility (as conceived by Anderson in terms of "responsive accountability"). By providing reasons to accept (or reject) *p* in the public forum, experts also expose themselves to the risk of being contradicted not only by other experts, but also by the lay citizens, who can in turn report the testimony of other experts as (weak, but depending on the inductive risk involved, also potentially strong enough) evidence.

Providing reasoned, rather than categorical, testimony allows citizens to bring different experts into dialogue, even if that dialogue never takes place officially. When expert A asserts "*p*, because *q*", and expert B asserts "*not-p* because *r*", lay people will have critical questions to pose both to A (*What about r? Is q coherent with not-p?*) and B (*What about q? Is r coherent with p?*). This creates new burdens of proof to be met, solicits additional reasons, creates further interactions among heterogeneous epistemic agents, and therefore allows the public discussion to move forward. Without reasoned testimonies, lay citizens would only be left with unqualified disagreement (*p* versus *not-p*), that does not allow for further progress in the discussion, creating a stalemate situation.

One way our criterion can be operationalized is then to require that expert testimonies have the form "*p*, because *q*". Mind that this does not require lay citizens to fully understand *p*, nor *q*, nor the causal link between *p* and *q*, as far as they can *use* them as testimonial evidence. This idea will then be in line with Feyerabend's suggestion to always *use* experts, in a way that accounts for lay people's role as legitimate interlocutors in public epistemic settings. It is by using them that experts can prove to be trustworthy and be, eventually, trusted.

6. *Conclusions*

Stemming from Feyerabend's cautionary stance on lay people's reliance and trust in experts, our examination of epistemic public trust offers a way toward

enhancing the dialogue between experts and the lay public. The argumentative dimension of public communication is not conceived here as an internal quarrel among disagreeing experts, nor is it merely about transferring information, but rather about cultivating a participatory dialogue where questions and criticisms are valued as essential components of knowledge development in the resolution of public problems. Part of the experts' trustworthiness will consist in recognizing and addressing through a public exchange of reasons the lay citizens' values, concerns, and knowledge – thus aligning scientific practices with societal needs and ethical standards. Thus understood, public communication settings can then allow trust and (mild) forms of skepticism to coexist productively.

Building upon Irzik and Kurtulmus' (2019) comprehensive proposal, we argue that to maintain and enhance epistemic public trust, experts must ensure they do not obstruct public discussions, especially when public interests are involved. This criterion can be operationalized by requiring experts to provide reasoned testimonies instead of categorical assertions. These testimonies will be then used as some kind of evidence by lay people, allowing public discussion to be carried on. Furthermore, adhering to this non-hampering criterion demonstrates experts' epistemic responsibility in a way that lay people can readily evaluate, thereby making it significantly relevant to public discourse. By demonstrating responsive accountability, experts can mitigate the risk of inappropriate epistemic dominance, aligning with Feyerabend's insights and concerns into the interplay between science and society. Echoing Feyerabend's advice, we suggest that trusting experts begins with the ability to effectively use them. However, unlike Feyerabend, we believe that this trust can be justified in meaningful ways.

Piero Avitabile
IMT School for Advanced Studies Lucca
Piero.avitabile@imtlucca.it

Alessandro Demichelis
IMT School for Advanced Studies Lucca
Alessandro.demichelis@imtlucca.it

References

Anderson, Elizabeth, 2011, "Democracy, Public Policy, and Lay Assessments of Scientific Testimony", in *Episteme*, 8: 144-164.

Baier, Annette, 1995, *Moral Prejudices: Essays on Ethics*, Harvard University Press, Cambridge (Mass.).

Ballantyne, Nathan, 2019, "Epistemic Trespassing", in *Mind*, 128: 510.

Barrotta, Pierluigi, Eleonora Montuschi, 2018, "The Dam Project: Who Are the Experts? A Philosophical Lesson from the Vajont Disaster", in Barrotta, Pierluigi, Giovanni Scarafile, eds., *Science and Democracy: Controversies and Conflicts*, John Benjamins, Amsterdam: 18-33.

—, Roberto Gronda, 2019, "Scientific Experts and Citizens' Trust: Where the Third Wave of Social Studies of Science Goes Wrong", in *Teoria*, 14: 9-27.

Besley, John C., Nicole M. Lee, Geah Pressgrove, 2021, "Reassessing the Variables Used to Measure Public Perceptions of Scientists", in *Science Communication*, 43: 3-32.

Brennan, Geoffrey, 2010, "The Division of Epistemic Labour", in *Analyse & Kritik*, 32: 231-246.

Collins, Harry, Robert Evans, 2007, *Rethinking Expertise*, The University of Chicago Press, Chicago.

Contessa, Gabriele, 2023, "It takes a Village to Trust Science: Towards a (Thoroughly) Social Approach to Public Trust in Science", in *Erkenntnis*, 88: 2941-2966.

Croce, Michel, 2019, "On what it Takes to be an Expert", in *The Philosophical Quarterly*, 69: 1-21.

Douglas, Heather, 2000, "Inductive risk and values in science", in *Philosophy of Science*, 67(4): 559-579.

—, 2009, *Science, Policy, and the Value-Free Ideal*, University of Pittsburgh Press, Pittsburgh.

Elliott, Kevin C., Daniel J. McKaughan, 2014, "Nonepistemic Values and the Multiple Goals of Science", in *Philosophy of Science*, 81(1): 1-21.

Fanelli, Daniele, 2010, "Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data", in *PloS one*, 5(4): e10271.

Feyerabend, Paul Karl, 1978, *Science in a Free Society*, New Left Books, London.

Gerken, Mikkel, 2022, *Scientific Testimony: Its Roles in Science and Society*, Oxford University Press, Oxford.

Goldman, Alvin I., 2001, "Experts: Which Ones Should You Trust?", in *Philosophy and Phenomenological Research*, 63: 85-110.

Hardwig, John, 1991, "The Role of Trust in Knowledge", in *The Journal of Philosophy*, 88: 693-708.

Irzik, Gürol, Faik Kurtulmus, 2019, "What Is Epistemic Public Trust in Science?", in *The British Journal for the Philosophy of Science*, 70: 1145-1166.

Jones, Karen, 1996, "Trust as an Affective Attitude", in *Ethics*, 107: 4-25.

Kitcher, Philip, 1990, "The Division of Cognitive Labor", in *The Journal of Philosophy*, 87: 5-22.

Lalumera, Elisabetta, 2024, "An Overview on Trust and Trustworthiness: Individual and Institutional Dimensions", in *Philosophical Psychology*, 37: 1-17.

Latour, Bruno, Steve Woolgar, 2013, *Laboratory Life: The Construction of Scientific Facts*, Princeton University Press, Princeton.

Martini, Carlo, 2019, "The Epistemology of Expertise", in Fricker, Miranda, Peter J.

Graham, David Henderson, Nikolaj J.L.L. Pedersen, eds., *The Routledge Handbook of Social Epistemology*, Routledge, London-New York: 115-122.

McCraw, Benjamin W., 2015, "The Nature of Epistemic Trust", in *Social epistemology*, 29: 413-430.

Origgi, Gloria, 2004, "Is Trust an Epistemological Notion?", in *Episteme*, 1: 61-72.

Reiss, Julian, Jan Sprenger, "Scientific Objectivity", in *The Stanford Encyclopedia of Philosophy (Winter 2020 Edition)*, Edward N. Zalta, ed., accessed February, 2024, https://plato.stanford.edu/archives/win2020/entries/scientific-objectivity/.

Resnik, David B., 2011, "Scientific Research and the Public Trust", in *Science and Engineering Ethics*, 17: 399-409.

Rudner, Richard, 1953, "The Scientist *Qua* Scientist Makes Value Judgments", in *Philosophy of science*, 20: 1-6.

Short, Daniel B., 2013, "The Public Understanding of Science: 30 Years of the Bodmer Report", in *School Science Review*, 95, 350: 39-44.

Solomon, Miriam, 2015, *Making Medical Knowledge*, Oxford University Press, Oxford.

Sosa, Ernest, 2011, "The Epistemology of Disagreement", in Haddok, Adrian, Alan Millar, Duncan Pritchard, eds., *Social Epistemology*, Oxford University Press, Oxford: 278-297.

Smolkin, Doran, 2008, "Puzzles about Trust", in *The Southern Journal of Philosophy*, 46: 431-449.

van Eemeren, Frans H., Rob Grootendorst, 2004, *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*, Cambridge University Press, Cambridge.

Walton, Douglas, 2010, *Appeal to Expert Opinion: Arguments from Authority*, Penn State Press, University Park.

Watson, Jamie C., 2020, *Expertise: A Philosophical Introduction*, Bloomsbury Publishing, New York.

Wilholt, Torsten, 2013, "Epistemic Trust in Science", in *The British Journal for the Philosophy of Science*, 64: 233-253.

Wynne, Brian, 1996, "May the Sheep Safely Graze? A Reflexive View of the Expert-Lay Knowledge Divide", in Lash, Scott, Bronislaw Szerszynski, Brian Wynne, eds., *Risk, Environment and Modernity. Towards a New Ecology*, Sage, London: 44-83.

Zagzebski, Linda T., 1996, *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*, Cambridge University Press, Cambridge.

—, 2012, *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*, Oxford University Press, Oxford.

Edizioni ETS
Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com - www.edizioniets.com
Finito di stampare nel mese di novembre 2024