

ISSN 1897-1652
e-ISSN 2154-3747

POLISH JOURNAL OF PHILOSOPHY

Volume XI, No. 1 (2017)

Institute of Philosophy
Jagiellonian University
Kraków

Published by the Institute of Philosophy of the Jagiellonian
University



POLISH JOURNAL OF PHILOSOPHY

SPRING 2017

Articles

- Valeria Bizzari
Aristotle, Phenomenology, and the Mind/Body Problem 7
- Douglas Ian Campbell
Against Lewis on 'Desire as Belief' 17
- Bogna Choińska
*Following Desire as the Ethical Postulate of
Psychoanalysis* 29
- Michael J. Shaffer
*Grounding Reichenbach's Pragmatic Vindication of
Induction* 43
- Adam Tamas Tuboly
*From 'Syntax' to 'Semantik' — Carnap's Inferentialism
and Its Prospects** 57

ARTICLES

Aristotle, Phenomenology, and the Mind/Body Problem

Valeria Bizzari

University of Pisa

Abstract. The mind-body relationship is a fundamental issue that has interested philosophers from very different schools of thought. Nowadays we can observe several positions being taken on this topic — my aim is to emphasize the phenomenological perspective on the mind-body relationship and, in particular, the role of Aristotelian thought in the contributions of philosophers such as Husserl and Merleau-Ponty. This paper consists of three different parts: in the first part, I will briefly sketch out a phenomenological account of the living body in Husserl and Merleau-Ponty; in second part, I will try to find parallels between phenomenology and Aristotle's philosophy. Finally, I will argue for an Aristotelian reading of schizophrenia, a pathology that seems to be caused by a disruption of the *corporeal Self*.

1. Husserl and Merleau-Ponty: the centrality of the *lived body*

One of the main achievements of the phenomenological method is undoubtedly that of revealing the *intentional* bond that connects the subject and the world: in other words, thanks to the concept of intentionality, Husserl seems to succeed in the attempt to explain the experience of the world in all its *concreteness*.

At first sight, this statement may sound like a paradox: the very act that “puts in brackets” the elements that fall under the label of „facticity” is the one that reveals the true essence of things. For this reason, it is possible to argue that the Husserlian subject has a *constitutive* role towards the world, and everything is related to it. This thesis is neither an idealistic argument — which would consider the person confined in a solipsistic position that transcends reality — nor a realistic one; rather, it tends to exceed both orientations in favor of a *correlation* between the world and consciousness that, as such, is unique and concrete: in other words, a temporal, intersubjective and *embodied* consciousness.

In contrast to a „dichotomous” vision of the Husserlian production and an approach that considers the body a prominent theme only in the „late” Husserl, as well as in opposition to an idealistic absolutist view, which assumes the Husserlian subject is a pure entity made up of synthetic capacities and which is only contingently related to the body, it seems to me that we can affirm not only the presence of the concept of „living body” in most of the production of the Master but also the centrality that this concept assumes for the most important aspects of the life of consciousness.

In fact, self-affection, which appears as reducible to the self-manifestation of subjectivity, is linked to two main functions: temporality—which provides the *formal structure* to the pre-reflective self-awareness, enabling the experience of my own lived acts as a unity made up of retentions, protections and living present—and the *embodiment*. In other words, perception implies a lived body as an orientation center provided with kinesthetic sensations. For instance, when I see or touch the surface of an object, I also co-perceive it through the kinesthetic sensations of the movements of my fingers. This is referring to an implicit self-awareness of my own body and of my own motor possibilities.

Being a *lived body* allows not only for self-consciousness, but also for the experience of the alterity. Firstly, because others are implied in my perceptual horizon before a concrete face-to-face encounter: perception is never exhausted in the short term, but is an original reference structure for potential and anonymous co-perceivers. Since this implicit co-perceiver is characterized by perception capabilities, he is necessarily embodied.

Furthermore, through the *Leib*, otherness is constituted as a new form of reality recognized as *analogous* thanks to the similarities between my body and that of others. Being embodied is essential: only through the *Leib* can I, in fact, enter into the field of perception of the other, which, in turn, recognizes me as *Leibanalogon*, a corporeality which finds its expression in an individuality.

The centrality of the living body is also present in the work of another important phenomenologist: Merleau-Ponty. The relevance that Husserl's thought had in the development of his philosophy emerges clearly in the preface to *Phenomenology of Perception* (Merleau-Ponty, 1945), his second doctoral thesis published in 1945, where the reference to the German phenomenologist is constant. The purpose of the text is the ontological rehabilitation of perception and of a corporeality conceived as a perceptual consciousness intentioning the world before any logical and representative operations.

Regarding specifically the theme of the body, it becomes clear how the Husserlian description of corporeality is the necessary starting point for the French phenomenologist: Husserl, in fact, stressed the importance of the mobility of the body, so necessary that there can be no perception without it, and emphasized the duality of the body in the subject's ontology (s/he is a material object, but also a living corporeity) and the consequent reversibility (consider, for example, the phenomenon of bilateralism of kinesthesia, where the subject is both „perceiver” and „touched”). Merleau-Ponty develops these themes, especially those concerning the essential involvement of the subject in the world: according to this view, which emerges already in Husserl's recent works, every experience is embodied, and the body appears as the middle term between Subjectivity and

Objectivity, irreducible to one or the other dimension because it is ontologically and essentially twofold. The question of the „primacy of perception,” the main theme of Merleau-Pontian philosophy, is entirely derived from the philosophy of Husserl, who — especially during the course „*The basic problems of Phenomenology*,” held in Göttingen from 1910 to 1911 — described sensory perception as a body issue because not only is perception itself intended as a purely physical talent, but the *body* itself is the „subject of perception.” Merleau-Ponty develops such theses and describes the body as the epistemological foundation of knowledge.

In this view, it seems to be possible to speak of the existence of a „phenomenology of the body” where the perceiver is a concrete subject, a living body, and understanding the role of the living body makes it possible to realize the essential complexity of existence.

2. Aristotle and the mind/body problem

Despite the fact that ancient philosophy seems to be strongly convinced of the separation between the mind and the body¹ (conceiving the mind as soul, as the spiritual substratum of the person), Aristotle represents an exception, and his interpretation of the link between the soul and its corporeal life is a source of complex and varying theories. On the one hand, there is the essential conception of the mind/body link,² according to which the soul is the essence of the body; on the other hand, instrumental and functionalist views consider the link between the soul and the body to be non-essential,³ because the body is a mere instrument or function.

According to my proposal, Aristotle’s ontology of the body looks very similar to the phenomenological one, because he emphasizes the role of the body both in perception and in the definition of the subject, who seems to be a “living corporeity.” Both phenomenology and Aristotelian philosophy, through the concept of *Leib* and the notion of *soma empsiuchon* respectively, led to the overcoming of a dualistic vision of the living being. But how did Aristotle face the mind/body problem?

¹ The separation between the mind and the body has been theorized upon since Homer, who put emphasis on the role that the soul — understood as a mere transcendent principle — had in defining the individual. Plato, in particular, set this trend by conceiving the soul as the principle of life, which is not only separated from the body, but, since it is immortal, detaches completely from the latter after its death. This approach has had remarkable fortune within the Christian tradition, favoring the emergence of a dualistic vision which found fertile ground both in the Middle Ages (for instance, in Thomas Aquinas’ theory, according to which the soul has no need for the body) and in the modern era.

² Cfr., for instance, F. Nuyens (1948), D. Ross (1924).

³ We can recall A. Bos and S. Menn among the instrumentalists, and Martha C. Nussbaum among the functionalists. Another position is a mere dualism, claimed, for instance, by R. Heinaman (see Heinaman, 1990).

In order to answer this question, I will mainly take into account the texts of *De Anima*, *Parva Naturalia*, and *Metaphysics* (book Z), and I will claim that in Aristotle's philosophy there is an essential link between the body and the soul, the same link that we can find in the phenomenological perspective.

Especially in Book II of *De Anima*, Aristotle reflects upon the ontological value of the soul. The *psiuchè* is described as the constitutive principle of the body: our soul cannot exist separately from the body (contrary to what Plato argued); rather, they are two unified elements which give rise to the living reality. To quote Aristotle himself: "the soul is not a body but requires a body; for it is not a body, but it belongs to a body, and for that reason it is present in a body, and in this sort of body" (414a20).⁴

In this view, the mind-body problem could be solved by linking the living body (the *Leib*, in phenomenological terms) to the psyche. The notion of *Leib* may find a parallel in the Aristotelian hylomorphic theory, which argues for the inextricable relation between our psychic capacities and the organic stratum (*sinolus*). Both the *Leib* and the *soma empsiucion* have the task of connecting the psychic functions with a material substrate, according to a gestaltic vision of reality, an *òlon*, where the different components take part in the whole. This does not mean arguing for a psychological materialism: both in Aristotle's philosophy and in phenomenology, the immateriality of the soul and its necessary link with the body stand together. In the first case, hylomorphism claims the inextricability between the soul and the body which it makes alive; in phenomenology, the notion of *Leib* is synonymous with the double ontology of the body, which is material *and* psychic, object and subject. The notions of *Leib* and that of *sinolus* (an entity composed of soul and body) are consistent with the thesis according to which the body allows the "actualization" of life. Following this perspective, we can argue that in Aristotle's philosophy we can find a vision of the living being as a *biological unity*, where the soul fulfills a large number of its functions *through* the body. According to Aristotle, the form realizes its own *ergon* precisely because it is an embodied principle, inextricably entangled in matter. The formal principle (*eidos*) linked to the organic substance ensures either the *specificity* of the sensible substance or the *universality* of the species to which it belongs.

Husserl's philosophy seems to be coherent with hylomorphism, especially in the late works of the father of phenomenology, where he underlines the existence of an essentially psychophysical subject. In *Ideas II*, in the chapter devoted to the analysis of the process of empathy, we can find an explicit reference to *De Anima*. As argued by Aristotle, according to Husserl the psyche is moving with the living body because *they are constantly entangled*. Husserl uses an Aristotelian passage in order to explain

⁴ The translation is mine, from the Italian edition of the second book of *De Anima*, translated into Italian by Giovanna R. Giardina.

the importance of movement in the intersubjective understanding: as the living body is a psycho-physical being characterized by a sensory-motor structure, so the *sinolus* is a sensible substance equipped with an internal motor principle. According to Husserl, seeing a moving body allows for the analogical argument and, subsequently, the recognition of a psycho-physical subjectivity, a *Leib* like me. In this perspective, the intersubjective process is also embodied.

Another similarity between Aristotle's doctrine and phenomenology is the importance given to *touch*. This specific sense is considered to be deeply important: according to Husserl; the tactile sensations of localization (which we can observe, for instance, in the phenomenon of bilaterality of kinesthesia) are the fundamental factor that leads to the synthetic constitution of the lived body, allowing subjective awareness. Although in different ways, the role of touch is emphasized also in Aristotle: this sense is considered the ground for all the others, and it is identified with the nourishment faculty, as if even intelligence would depend upon sensitivity. On this view, perception seems to be the ground for intellectual apprehension: the science of soul is in continuity with biological and physical inquiries, because there is a strong link between intellectual activity and corporeity.

Either in phenomenology (especially in Merleau-Ponty's thought, whose emphasis on the role of the body is stronger than in Husserl's thought) or in Aristotle's philosophy, corporeity is revalued *epistemologically* and *ontologically*. From an epistemological point of view, in the *Metaphysics* and in the *Posterior Analytics*, the descriptions of sensorial activity (intended as *esthesis*, simple sensation, and as *empeiria*, empirical experience) are consistent with attributing a gnoseological value to sensitivity, which seems to be the ground for more complex intellectual activities (like imagination or memory). In the same way, Husserl underlined the importance of kinesthesia and of passive synthesis, and Merleau-Ponty argued for "the primacy of perception"; in both authors, the relationship between the subject and the world seems to resolve itself in a tangle of cognition and action. From this point of view, the world is always perceived from a specific perspective of which my living body is the center. Therefore, objects have a specific meaning depending on the potentialities of action of the subject, which have been intuitively grasped and are not the result of inferential processes. This phenomenological thesis seems to be an echo of Aristotle's philosophy, also in emphasizing the importance of movement:⁵ we must remember, in fact, that the Stagirite affirmed: "We must therefore see that we understand what

⁵ Nowadays, Aristotle is used as a reference by some phenomenologists who strongly emphasize the role of bodily movement in perception and cognition (a paradigmatic case is Maxine Sheets-Johnstone (see, for instance, Sheets-Johnstone (2011)).

motion is; for if it were unknown, nature too would be unknown.” (Aristotle, *Physics* 200b12-14)⁶

From an ontological perspective, while in phenomenology we can talk about a *Leib*, an embodied subjectivity linked with the world in a pre-reflexive and intentional manner, in Aristotle the *soma empsiucon* is the result of the indissoluble union between psyche and matter, and the soul itself seems to be (as we can notice in *De Anima*) “the first act of a natural body with the capacity for life” (412a27, my translation from Italian).

A concrete example of the entanglement between the mind and the body is represented by the description of the dreamlike experience in *De Somno*: while, on the one hand, Aristotle claims that during this process the subject can make up for those elements which are latent when he/she is awake (bringing forward Freud’s intuition), on the other hand he gives us a detailed account of dreams as psycho-physical phenomena which imply anatomical and physiological processes. For instance, he describes the dream as an experience that depends on the condition of the heart, of the sensorial organs, even on the pureness of the bloodstream. It is very interesting to notice that neuroscientific discoveries have confirmed, in a certain sense, Aristotle’s intuition: in fact, nowadays we know that our hippocampus adopts different modes of operation when we are sleeping. We can then affirm that Aristotle (and later, phenomenology) seems to be right in conceiving the subject as ontologically and epistemologically bound with corporeity.

But what are the consequences of a similar (embodied) approach in the contemporary debate? I will try to answer to this question in the next section, where, using Aristotelian and phenomenological vocabulary, I will focus on a case study of (disrupted) perception.

3. Aristotle and the role of the body in perception: a case study

We can practically observe this embodied perspective “at work” in the analysis of perception, especially in those cases where perception is disrupted. In fact, it is precisely on these occasions that the body reveals its importance. In phenomenological psychopathology, there is a strong tendency to consider corporeal awareness to be the very ground of synthetic capabilities and consciousness activities.⁷ In fact, our perception and our “being in the world” are usually permeated by a bodily sense of Self: recent empirical studies have shown that this *corporeal* feeling is present from the

⁶ Barnes (1984).

⁷ See, in particular, the works by Stanghellini (especially 2006) Fuchs (2005) and Fuchs, Schlimme (2009).

first years of our life, like a primordial “sensus communis” that is the ground for abilities like proprioception, perception, and emotion.⁸

The definition of common sense (or *koinè aisthesis*) is provided by Aristotle, who claims: “Besides the specific sense there is the *sensus communis*, which is not a sixth sense but a generic power of sensation as such which *provides unity for the sensitive soul* in its particular manifestations. The ear does not see; however, the man who hears also sees, and some qualities are presented through more than one sense... We also perceive that we are perceiving through *sensus communis*.”⁹

My thesis is that *koinè aisthesis* is an *embodied* act: in fact, *providing unity for the sensitive soul* seems to be an activity very similar to the Husserlian passive synthesis, where corporeity plays a fundamental role. Common sense allows the subject to combine the different sensorial modalities, associating each sensation with a form of self-awareness. As a confirmation, in *Nicomachean Ethics* Aristotle affirms that the awareness of perception itself means the existence of perception. For this reason, we can say that self-awareness is sensorial and embodied. In the same way, we have seen how in Husserlian philosophy there is a pre-reflexive, passive, and anonymous self-consciousness, antecedent to any synthetic activity and presupposing a lived body which works as an orientation center. In fact, kinesthetic sensations make us aware of our sensations and movements immediately and intuitively, in a sort of primary self-consciousness.

In schizophrenic subjects we can register a progressive alienation from their own bodily feeling. The disembodiment of the self is often the culprit of the loss of perceptual and cognitive capacities: in this perspective, our lived body could be considered the core of our tacit and pre-reflective understanding of the environment, while its loss necessarily seems to involve a loss in the sphere of *koinè aisthesis*.

Furthermore, as G. Stanghellini noticed (2004), the supramodal perception described by the Stagirite, i.e. the pre-reflexive and pre-conceptual sensorial synthesis, seems also to be the ground for the understanding of alterity. In other words, intersubjectivity can be defined in the same way as intercorporeity, as a process based on the immediate transfer of corporeal schema. The consequence is that, if the embodied being of a subject is compromised, his self-consciousness and his ability to find attunement with the other and the world will be lost or disrupted.

⁸ I mean, for instance, the studies by Meltzoff and Moore (Meltzoff, Moore, 1977; or Moll, Meltzoff, 2010). According to Meltzoff, a *corporeal* attunement is the constitutive ground for social understanding, from its first step (sharing perspective) to the more complex level of understanding. In other words, our pre-reflective and corporeal sense of self is immediately linked to intersubjectivity and it allows us to have both proprioception and the perception of alterity.

⁹ Edward (1967, p.3).

Social and corporeal selves seem to share the same roots, constituted by the sensory-motor integration given by *koinè aisthesis*. In the case of schizophrenia, the importance of similar grounding is very clear: in fact, the subject is affected by *kinesthetic disorders*. The senses are perceived as fragmented, and the subject experiences a disruption of self-awareness, which is one of the main symptoms preceding the detachment of the self from intersubjective world.

As I have previously noted (in the first section), in Husserlian philosophy the subject is intentionally linked to the world through his kinesthetic capabilities, originarily connected to his living corporeity. In the same way, it is possible to claim that the sensory-motor organization described by Aristotle is the necessary condition for an adequate perception of alterity. The sensorial coordination at the base of this mechanism is the same structure that is disrupted in the schizophrenic experience, especially in the first stages of development of this pathology: in fact, we can register abnormal bodily sensations which cause disorders in the intercorporeal attunement process at the root of intersubjectivity. What emerges is the particular importance of a sensorial and kinesthetic consciousness, which seems to be consistent both with Aristotle's thought and a phenomenological analysis.

Conclusion

In both phenomenology and Aristotle's philosophy, the role of the corporeal structure of self seems to be fundamental: it is necessary in order to understand the world and in order to have social attunement. More specifically, the link between the mind and the body comes to light in Aristotle's analysis of perception, where we can find the gnoseological value of sensation that looks to be coherent with a "phenomenology of embodiment." The mind and the body, form and matter, give rise to a biological unity: on this view, the mind/body problem seems to be solved, or we can even say that it is a relationship which does not need to be considered a problem at all. We can simply argue that the soul is form and the body is matter. Since they constitute a *sinolus*, mind and body, soul and matter cannot be separated. In the same way, from a phenomenological perspective, *Leib* and *Körper* are synonymous with the double ontology of the person, which is an embodied consciousness. To put it more appropriately: "Hence we need not ask whether the soul and body are one, any more than we need to ask this about the wax and the seal or, in general, about the matter and the thing of which it is the matter. For while one and

being one are spoken of in several ways, the actuality <and what it actualizes> are fully one.”¹⁰

References

- Aristotle (2009). *De Anima (Second Book)*. Italian Transl. by G. Giardina, Roma: Aracne.
- Barnes, J.(ed.) (1984). *The Complete Works of Aristotle*, Vol. 1 (pp. 315-446) NJ: Princeton University Press.
- Bos, A. (2003). *The Soul and its Instrumental Body*. Leiden-Boston: Brill.
- Edward, P.(ed.) (1967). *The Encyclopedia of Philosophy*, vol. 7. New York: Collier Macmillan.
- Fuchs, T. (2005). Corporealized and Disembodied Minds: a Phenomenological View of the Body in Melancholia and Schizophrenia. *Philosophy, Psychiatry and Psychology*, 12, 95-107.
- Fuchs, T., Schlimme, J. E. (2009). Embodiment and Psychopathology: a Phenomenological Perspective. *Current Opinion in Psychiatry*, 22, 570-575.
- Heinaman, R. (1990). Aristotle and the Mind/Body Problem. *Phronesis*, 35(1), 83-102.
- Husserl, E. (1952). *Ideen zu einer reinen phänomenologischen Philosophie. Zweites Buch*. Martinus Nijhoff: Den Haag.
- Meltzoff, A., Moore, K. M. (1977). Imitation of Facial and Manual Gestures by Human Neonates. *Science*, 198, 75-78.
- Menn, S. (2002). Aristotle’s Definition of Soul and the Programme of the *De Anima*. *Oxford Studies in Ancient Philosophy*, 22, 83-139.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Paris: Gallimard.
- Moll, M., Meltzoff, A. N. (2011). Perspective Taking and Its Foundation in Joint Attention. in J. Roessler(ed.) *Perception, Causation, and Objectivity. Issues in Philosophy and Psychology* (pp. 286-304). Oxford: Oxford University Press.
- Nuyens, F. (1948). *L’évolution de la psychologie d’Aristotle*. Louvain, Éditions de l’Institut supérieur de philosophie; La Haye, Martinus Nijhoffs Uitgevers: Paris, Vrin.
- Nussbaum, M. C., Rorty, A. (2003). *Aristotle’s Essay on De Anima*. Oxford Scholarship Online.
- Ross, W. D. (2003). *Aristotle*. London: Routledge.
- Sheets-Johnson, M. (2011). The Corporeal Turn. *Journal of Consciousness Studies*, 18 (7-8), 145-68.
- Stanghellini, G. (2006). *Disembodied Spirit and Deanimated Bodies*, Oxford: Oxford University Press.

¹⁰ Aristotle (2009, 412b6-9).

Against Lewis on ‘Desire as Belief’

Douglas Ian Campbell
The University of Canterbury

Abstract. David Lewis describes, then attempts to refute, a simple anti-Humean theory of desire he calls ‘Desire as Belief’. Lewis’ critics generally accept that his argument is sound and focus instead on trying to show that its implications are less severe than appearances suggest. In this paper I argue that Lewis’ argument is unsound. I show that it rests on an essential assumption that can be straightforwardly proven false using ideas and principles to which Lewis is himself committed.

1. Introduction

David Lewis (Lewis, 1988, 1996) describes, then attempts to refute, an anti-Humean theory of desire he calls ‘Desire as Belief’ (DAB). I will call Lewis’ argument against DAB the ‘updating argument’. The implications of the updating argument are often held to extend far beyond disputes over the nature of desire. For example, Graham Oddie says it constitutes ‘a massive problem for realism about value’ (Oddie, 1994, p. 453), while Ruth Weintraub says it ‘entails subjectivism about ethics’ (Weintraub, 2007, p. 119).

In this paper I show that the updating argument rests on an essential assumption that can be straightforwardly proven false using ideas and principles to which Lewis is explicitly committed and which are central to the updating argument itself. My position, in other words, is that in constructing the updating argument Lewis implicitly contradicts himself.

The error I identify in the updating argument appears to have gone unnoticed by the updating argument’s many other critics. These critics have, with a few exceptions, accepted the updating argument as sound, and focused instead on trying to show that its soundness can be tolerated. A common tack has been to contend that DAB is not in any case a plausible or attractive (or even, perhaps, coherent) version of anti-Humeanism, and then argue (or at least suggest) that certain more plausible versions of anti-Humeanism are invulnerable to the updating argument (Broome, 1991; Byrne & Hájek, 1997; Daskal, 2010; Hájek & Pettit, 2004; Price, 1989; Weintraub, 2007). This approach has the weakness that, while it might provide comfort to the anti-Humean, it does nothing to dispel the threats posed by the updating argument to value-realism and objectivism in ethics. The present paper disposes of all such threats.

§2 describes DAB. §3 outlines the updating argument. §4 exposes the error in the updating argument. §5 deals with possible Lewisian rejoinders. §6 wraps things up.

2. DAB

Following Lewis (Lewis, 1988, 1996), let ' A° ' (pronounced ' A halo') denote the claim *that A 's being true is good*. Harmlessly simplifying (although c.f. (Oddie, 2001)), Lewis ignores *degrees* of goodness, and in the interests of generality he doesn't specify what A 's being good consists in (instead leaving one free to plug in whichever analysis of goodness one wants). Let $V(A)$ denote the degree to which a rational agent values A being the case. Let $C(p)$ denote a rational agent's credence for p 's being true. DAB is the following claim:

$$\text{DAB:} \quad V(A) = C(A^\circ)$$

That is, DAB is simply the idea that the value assigned by a rational agent to some proposition A , and thus the extent to which she desires that A , is identical to her credence for A being good.¹ DAB is 'anti-Humean' in the sense that it entails a rational agent's desires logically supervene on her beliefs about what is good: i.e., that her desires and beliefs are not modally separable 'distinct existences'.

In the above formula, A is intended by Lewis to be a universally quantified variable ranging over all propositions. Following Lewis, I leave this implicit, omitting the quantifier. For simplicity I continue to suppress quantifiers until near the end of the paper.

3. Lewis' updating argument

The updating argument (Lewis, 1988, 1996) involves the following two claims:

$$\begin{array}{ll} \text{INV:} & V(A|A) = V(A) \\ \text{X:} & V(A|A) = C(A^\circ|A) \end{array}$$

INV says, in effect, that the value a rational agent assigns to a proposition, A , being true will not be affected by her learning that A is indeed true. Lewis provides a detailed argument for INV (Lewis, 1988, pp. 331-332). Recently, several authors have argued that INV is false (Bradley & List,

¹ Although DAB is intended as a theory of *desire*, Lewis casts it in terms of the V operator, and thus in terms of what the agent *values*. His tacit assumption is that an agent desires a given state of affairs just to the degree she values it. If this assumption appears dubious then it can simply be stipulated that $V(A)$ denotes strength of desire, rather than level of valuing.

2009; Bradley & Stefánsson, 2016; Stefánsson, 2014). But here I will assume that INV is true, and show that there is a problem elsewhere in Lewis' argument.

X is clearly a claim that a proponent of DAB must accept: for if a rational agent has learnt that A obtains, then her credence for A being good, and thus — according to DAB — the degree to which she values A , will be updated by conditionalization on A . This is what X says.

DAB, INV and X together entail IND:

$$\text{IND: } C(A \upharpoonright A) = C(A \circ)$$

Thus, a proponent of DAB must accept IND. However, suppose an agent's credence function is as depicted in the Venn diagram of Figure 1 (Lewis, 1996, p. 309).

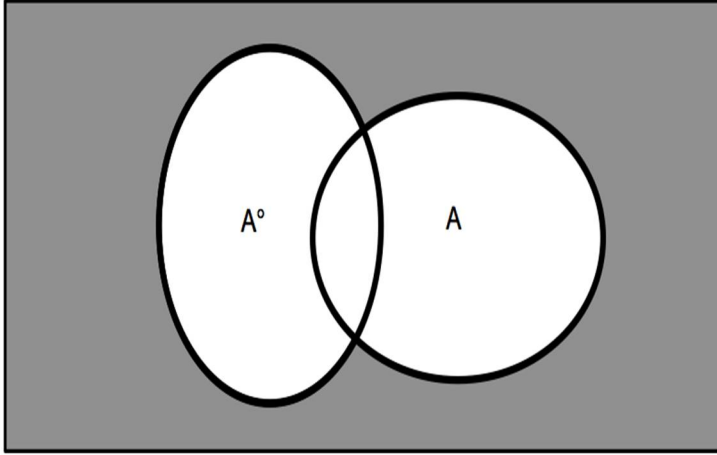


Figure 1. An agent's credence function for A , A° , and their negations.

Suppose that the agent learns (say) that $(A \vee A^\circ)$ is the case, and updates her credences accordingly. Let's call her updated credence function, C' . By the rules of Bayesian updating, C' is obtained by taking the credence represented by the shaded region of Figure 1 and redistributing it to the remaining regions so as to leave the relative sizes of these remaining regions unchanged. When this is done the A° -region grows in size. Thus $C'(A^\circ) > C(A^\circ)$. On the other hand, the proportion of the A -region which overlaps with the A° -region doesn't change. Thus $C'(A \upharpoonright A) = C(A \upharpoonright A)$. When these two results are combined with the assumption that C satisfies IND, we get:

$$C'(A \upharpoonright A) = C(A \upharpoonright A) = C(A^\circ) < C'(A^\circ),$$

and thus,

$$C(A \circ A) < C(A \circ).$$

Notice that this means the agent's updated credence function, C' , *does not conform to IND*. And so even if an agent starts off using a credence function which satisfies IND, this property — of satisfying IND — will be unstable under Bayesian updating. It will, for instance, be lost as soon as the agent learns that $(A \vee A \circ)$ is the case and updates her credences accordingly. A proponent of DAB must — Lewis says — therefore deny that credences are to be adjusted by Bayesian updating. However Lewis insists, reasonably enough, that given the choice of repudiating DAB or repudiating Bayesian updating, we should repudiate the former, not the latter (Lewis, 1988, p. 325).

4. The error in Lewis' reasoning

My refutation of Lewis' updating argument has two steps. First, I prove that a result I call the 'conditionalization conjecture' (CC) must be true if DAB is true. Next I show that if CC is true, then Lewis' updating argument is unsound.

Suppose that $A \wedge A \circ$ is true — which is to say that A is true, and that A is good. Then it follows, obviously, *that a good state of affairs obtains*. Let this proposition, that a good state of affairs obtains, be called H (for 'halo'). This being so, $H \circ$ says it would be good that a good state of affairs obtains, which is a trivial truth. On the other hand, $(\neg H) \circ$ says it would be good that a good state of affairs does not obtain, which is a trivial falsehood. A rational agent will therefore assign a credence of 1 to $H \circ$, and a credence of 0 to $(\neg H) \circ$. These results are recorded as follows:

$$C(H \circ) = 1 \tag{1}$$

$$C((\neg H) \circ) = 0 \tag{2}$$

If we assume DAB, then from (1) and (2) we can derive (3) and (4):

$$V(H) = 1 \tag{3}$$

$$V(\neg H) = 0 \tag{4}$$

According to Lewis (Lewis, 1988, p. 326, 1996, p. 303), the values assigned to outcomes by a rational agent will obey a principle of additivity, which he characterises as follows: 'the value of a proposition that might come true in several alternative ways is an average of the values of those several alternatives, weighted by their conditional credences' (Lewis, 1988, p. 326). In symbols:

$$V(X) = \sum_i (V(X \wedge E_i) C(E_i | X)) \tag{5}$$

Here X denotes any given proposition, and $\{E_1, \dots\}$ denotes any given set of mutually exclusive and jointly exhaustive propositions (i.e., any 'partition' of possibility space).

Now, let T represent some obvious tautology (say, ' $1=1$ '). Consider the partition $\{A, \neg A\}$, where A is some arbitrary proposition. Substituting T for X and $\{A, \neg A\}$ for $\{E_1, \dots\}$ within (5) gives (6):

$$V(T) = V(T \wedge A)C(A|T) + V(T \wedge \neg A)C(\neg A|T) \quad (6)$$

When (6) is simplified in light of T being a tautology (and hence a necessary truth), it turns into (7):

$$V(T) = V(A)C(A) + V(\neg A)C(\neg A) \quad (7)$$

Now, let's suppose the arbitrary selected proposition, A , that is mentioned in (7) is in fact identical to H . (7) then becomes (8):

$$V(T) = V(H)C(H) + V(\neg H)C(\neg H) \quad (8)$$

Substituting (3) and (4) into (8) yields (9):

$$V(T) = C(H) \quad (9)$$

Substituting (9) back into (7) gives us (10):

$$C(H) = V(A)C(A) + V(\neg A)C(\neg A) \quad (9)$$

Conditionalizing on A (and thus setting $C(A)=1$ and $C(\neg A)=0$) we obtain (10):

$$C(H|A) = V(A) \quad (10)$$

Assuming again that DAB is true, we can derive CC from (10):

$$\text{CC: } C(A^\circ) = C(H|A)$$

CC says that, for any proposition A , the credence that an agent assigns to A° is identical to the *conditional credence* she assigns to *there being a good outcome given that A comes true*. Notice that CC implies that $C(A^\circ)$ is not an *unconditional* credence (as surface appearances suggest), but that it is instead a *disguised conditional credence* – namely, the conditional credence of H given A .

So much for why, if DAB is true, CC must be true. I now explain why, if CC is true, then Lewis's updating argument against DAB rests on an illegitimate assumption.

As already explained, the updating argument rests on the assumption that credences are distributed as in Figure 1, with one 'dollop' of credence being assigned to A , another, overlapping 'dollop' of credence being assigned to A° , and the conditional credence assigned to $(A^\circ|A)$ being represented by the proportion of the A -region which overlaps with the

A° -region. If credences are distributed in this way, then Lewis is quite right: Bayesian updating can in this case potentially change how much credence is assigned to A° without causing any corresponding change to how much conditional credence is assigned to $(A \circlearrowleft A)$ (or *vice versa*) – with the result that IND will not be robustly satisfied.

However, if CC is true, then Figure 1 misrepresents the logical relationship between A and A° . Whereas Figure 1 represents A and A° as being logically independent propositions, CC implies that they are not independent, and that A° is instead a conditional proposition, conditioned on A . If CC is true, then a rational agent's credences must be distributed not as they are in Figure 1, but rather as they are in Figure 2:

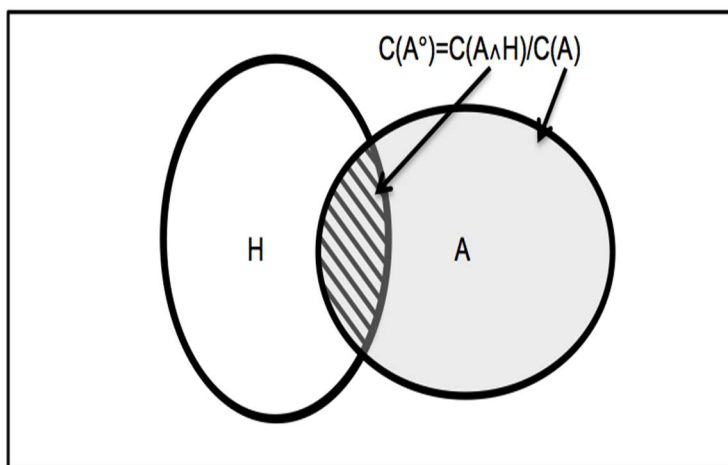


Figure 2. How credences should be assigned by a rational agent.

In Figure 2 one ‘dollop’ of credence is assigned to H , and another to A . The amount of credence assigned to A° , $C(A^\circ)$, is identified with $C(H|A)$, and thus with the proportion of A -region that overlaps with the H -region – or, that is, with the proportion of the shaded region that is crosshatched.

Now, notice two points. First, under this way of allocating credences, $C(A^\circ)$ is a conditional credence, and it is conditioned on A (for it is the conditional credence of H given A). Second, it is a universal rule that if a credence function is already conditioned on a certain outcome, then updating it by conditionalizing it on the same outcome again has no effect. (That is, for any two credence functions P and Q , if $P(x) = Q(x|a)$ – i.e., if P is already conditioned on a – then $P(x|a) = P(x)$ – i.e., conditionalizing P on a again changes nothing.) Putting these two points together, it follows that $C(A \circlearrowleft A) = C(A^\circ)$ (for, since A° is already conditioned on A , conditioning on A again has no effect). Thus, it follows that $C(A \circlearrowleft A)$ will be represented in

Figure 2 in exactly the same way that $C(A^\circ)$ is represented – namely, by the proportion of the shaded region that is crosshatched. Since $C(A^\circ)$ and $C(A^\circ|A)$ are, for this reason, necessarily identical, it is impossible for Bayesian updating to produce any mismatch between them. That is, not only are $C(A^\circ)$ and $C(A^\circ|A)$ both represented in Figure 2 by the proportion of the shaded region that is crosshatched, *but they will both continue to be represented by this single element of the diagram even as the various regions in Figure 2 grow and shrink when forces of Bayesian updating are operating upon them.* Since $C(A^\circ)$ and $C(A^\circ|A)$ must, for this reason, always be identical, IND will always be satisfied.

Let's take stock. Lewis' updating argument against DAB allegedly proves that no credence function can robustly satisfy IND in the face of Bayesian updating. However, the updating argument assumes that credences are distributed as per Figure 1. I have shown that if DAB is true, then CC is true, and that if CC is true, then credences must instead be distributed as per Figure 2. Under this way of distributing credences, IND is necessarily satisfied. And so, in arguing against DAB, Lewis begs the question against DAB, by making an assumption that is inconsistent with the truth of DAB. The updating argument is unsound for this reason.

5. How might Lewis respond?

CC tells us that for any credence function, C , and proposition, A , $C(A^\circ) = C(H|A)$. The implication is that the halo function, A° , is a propositional function that accepts a proposition, p , as input, and produces another proposition, p° , as output, such that a rational agent's credence for p° should always match her conditional credence for H , given p . But is a propositional function that meets these specifications even logically possible in the first place? We need look no further than some of David Lewis's own earlier work in order to find an argument that purports to show that such a function is *not* logically possible.² In his (Lewis, 1976, 1986) Lewis aims to demonstrate that there can be no systematic way of mapping a pair of propositions, q and r , onto a proposition, s , such that $C(s)$ should match $C(q|r)$. Lewis might respond to my demonstration that DAB entails CC by harking back to this earlier work. Specifically, he could argue that if DAB entails CC, then it follows from his earlier work that the halo function cannot exist, and thus that DAB is incoherent (because it posits the halo function).

Although Lewis could respond in this way, there are two points to notice. First, although this would still give Lewis an argument against DAB, it

² Many thanks to an anonymous reviewer for pointing me to this work of Lewis.

would be a different argument than the one he in fact uses, the updating argument, which I have shown to be unsound. Second, it is controversial whether Lewis' earlier results are correct (see, e.g., Edgington, 2014; Milne, 1997). Hence, they provide an insecure foundation for a revised argument against DAB.³

Another reply Lewis might make to my rebuttal of his updating argument can be guessed based on his response to a related argument by Huw Price. Price (1989) shows that Lewis' updating argument can't be used to refute the following anti-Humean rival to DAB, which Lewis (Lewis, 1996) calls 'Desire as Conditional Belief' (DACB):

$$\text{DACB:} \quad V(A) = C(A \mid A)$$

Lewis' (1996) response to Price is rather involved. Here is my reconstruction of it. Lewis begins by, in effect, distinguishing two claims to which the Humean is committed:

H1. Facts about what a rational agent desires are not fixed or determined by facts about what she believes.

H2. Agents that are alike in being rational, that have the same priors, and that are privy to the same empirical data about the world, can potentially be unlike in what they value and desire.

He then distinguishes two versions of anti-Humeanism, which I will call 'strong' and 'weak'.⁴ The strong anti-Humean denies both H1 and H2, while the weak anti-Humean denies only H1. Lewis represents the strong anti-Humean's claim as follows:

$$\text{DBN:} \quad V(A) = C(G \mid A)$$

Here G amounts to the proposition *that something objectively good happens* (Lewis, 1996, p. 307). Since DBN is intended to express strong anti-Humeanism, it should contradict both H1 and H2, and indeed it does. It contradicts H1 because it implies that facts about what a rational agent values, and thus desires, are fixed by her credences, and hence by her beliefs. It contradicts H2 because it implies that rational agents are all alike in having desires geared toward the same ultimate, objective good, embodied by G .

Lewis is surprisingly non-hostile to strong anti-Humeanism as encapsulated in DBN. He says little about it — only that it is

³ My own view is that Lewis's (1976) result is definitely erroneous, but I will defend this claim elsewhere.

⁴ Lewis (1996) instead calls these positions 'Desire by Necessity' and 'Desire as Belief'. However, these terms are ambiguous since he uses them *both* as names of the two anti-Humean positions in question *and* as the names of formula that purportedly express these anti-Humean positions. Moreover, the terms are not especially clear — so I don't use them.

a 'comparatively simple and unproblematic version' of anti-Humeanism (Lewis, 1996, p. 307). He makes no attempt to refute it. His lack of interest in refuting strong anti-Humeanism is presumably to be explained in terms of his being a sufficiently committed Humean to think strong anti-Humeanism has no real chance of being true (due to its radical denial of H2). Weak anti-Humeanism poses a much more credible threat to Humeanism.

Lewis proceeds to critique various possible formulations of weak anti-Humeanism, including both his own DAB and Price's DACB. As we have seen, he holds that DAB is untenable because it falls prey to the updating argument. His response to Price, in contrast, is that DACB turns out, upon analysis, *to be re-expressible as DBN*. His complaint against DACB is therefore that *it is not a version of weak anti-Humeanism after all*. It is just strong anti-Humeanism in disguise. In Lewis' words, DACB is 'a form of anti-Humeanism, sure enough, but not the right form of anti-Humeanism' (Lewis, 1996, p. 313).

Now, in my above rebuttal of the updating argument I showed that CC is true, and when DAB and CC are put together they entail a theory strikingly similar to DBN. (More on this shortly.) This being so, Lewis might reply to me much as he replies to Price. Specifically, he might contend that I have rescued DAB from the jaws of the updating argument only by showing that DAB amounts to a form, not of *weak* anti-Humeanism, but of DBN, and thus of *strong* anti-Humeanism. In other words, Lewis might, while conceding that I have refuted his updating argument (a major concession on his part), claim that I have nevertheless played into his hands by providing an alternative proof of the result he is after: viz., that DAB is not a tenable form of *weak* anti-Humeanism.

My counter-reply to this potential reply of Lewis' is to deny that DAB and CC together entail a form of strong anti-Humeanism. (I believe Price could respond similarly with regards DACB.)

DAB and CC together entail a theory I will call 'DAB⁺':

DAB:	$V(A)=C(A \text{ } ^\circ)$
CC:	$C(A \text{ } ^\circ)=C(H A)$
DAB ⁺ :	$V(A)=C(H A)$

DAB⁺ and DBN certainly *appear* to be very similar theories. Indeed, they appear to share precisely the same logical form, with the only difference between them being that where DAB⁺ references *H*, DBN instead references *G*. But this apparent similarity is, I will now show, merely superficial. When quantifiers are brought into the open, the two views are exposed as having *different logical forms*.

To see this, let's start with DAB. Lewis proposes DAB as a formulation of *weak*, rather than *strong*, anti-Humeanism. This raises a puzzle. On the face of it, since DAB says simply that $V(A)=C(A \circ)$, DAB would seem to imply that if two rational agents share all the same beliefs, and thus assign the same credences to all propositions, then they will necessarily share the same values, and so share the same desires. Thus, DAB would appear to be inconsistent with the possibility of a pair of equally well-informed rational agents with the same priors diverging in what they desire. If this were right then DAB would entail the falsity of H2 as well as the falsity of H1, and so it would be a form of *strong* anti-Humeanism, not of *weak* anti-Humeanism. The solution to this puzzle lies in recognizing that the halo function is intended by Lewis to be *agent-relative*. Different rational agents can potentially have different halo functions. That is, it can potentially be the case that for one agent, $A \circ = P$, while for a second agent, $A \circ = Q$, where P and Q are different propositions. If these two agents share all the same beliefs, then they will agree in the credences they assign to P and Q respectively. But they will nevertheless differ in how much they value, and thus desire, A 's being the case, for where the first agent is concerned, $V(A)=C(A \circ)=C(P)$, while for the second agent, $V(A)=C(A \circ)=C(Q)$.

Recall that H is characterized by reference to the halo function (since H is defined such that (1) and (2) come out as true). This being so, it follows from the fact that Lewis' halo function is agent-relative that H will be agent-relative too. With this in mind, consider the following two rival ways in which DAB⁺ might be formulated when quantifiers are brought into the open:

$$\text{DAB}^+1: \quad \exists H \forall x \forall A: V_x(A)=C_x(H|A)$$

$$\text{DAB}^+2: \quad \forall x \exists H \forall A: V_x(A)=C_x(H|A)$$

Here $V_x(A)$ denotes the value assigned by a rational agent, x , to an outcome, A . Similarly, C_x denotes the credence function used by x . Notice that, because of where the existential quantifier is positioned, DBN⁺1 implies that there is a single proposition, H , which serves as a universal 'yardstick' against which *every* rational agent measures the values of outcomes. On the other hand, DBN⁺2 is consistent with the possibility of different agents using different versions of H as their respective yardsticks for measuring value. Clearly the idea that the halo function, and thus the identity of H , can *vary* between rational agents is correctly captured by DBN⁺2, but not by DBN⁺1. We should therefore understand DBN⁺ as being equivalent to DBN⁺2, not as being equivalent to DBN⁺1.

Let's now turn to DBN, two rival formulations of which are as follows:

$$\text{DBN}1: \quad \exists G \forall x \forall A: V_x(A)=C_x(G|A)$$

$$\text{DBN}2: \quad \forall x \exists G \forall A: V_x(A)=C_x(G|A)$$

DBN is, of course, intended by Lewis to encapsulate *strong* anti-Humeanism, with the idea being that *G* represents an *objective* good, that is the same for all rational agents and cannot vary between rational agents. This being so, it is obvious that DBN is to be correctly understood as saying the same thing as DBN1, not DBN2.

Thus, when quantifiers are brought into the open DAB^+ is revealed as being equivalent to DAB^+2 , while DBN is revealed as being equivalent to DBN1. Crucially, DAB^+2 and DBN1 *do not* share the same logical form. Hence, initial appearance to the contrary notwithstanding, DAB^+ and DBN do not share the same logical form either. They *appear* to share the same logical form only because of a scope ambiguity. When this ambiguity is resolved by bringing quantifiers into the open, it becomes clear that DAB^+ is consistent with rational agents *differing* among each other in what they ultimately value, as *per* weak anti-Humeanism, while DBN instead entails that rational agents must be *alike* in what they ultimately value, as *per* strong anti-Humeanism. The potential objection here being considered — that by showing DAB entails DAB^+ I have shown DAB to be a form of strong anti-Humeanism — is thus without foundation.

6. Conclusion

Lewis' updating argument is, I believe, decisively refuted by the above proof of CC. If this is right then anti-Humeans (weak anti-Humeans included), value-realists, and objectivists about ethics have nothing to fear, at least where this particular threat to their philosophies is concerned.

References

- Bradley, R., List, C. (2009). Desire-as-Belief Revisited. *Analysis*, 69(1), 31-37.
- Bradley, R., Stefánsson, H. O. (2016). Desire, Expectation, and Invariance. *Mind*, 125(499).
- Broome, J. (1991). Desire, Belief, and Expectation. *Mind*, 100(2), 265-267.
- Byrne, A., Hájek, A. (1997). David Hume, David Lewis, and Decision Theory. *Mind*, 106(423), 411-428.
- Daskal, S. (2010). Absolute Value as Belief. *Philosophical Studies*, 148(2), 221-229.
- de Finetti, B. (1995). The Logic of Probability. *Philosophical Studies*, 77(1), 181-190.
- Edgington, D. (2014). Indicative Conditionals. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy* (Winter 2014). Metaphysics

- Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/win2014/entries/conditionals/>
- Hájek, A., Pettit, P. (2004). Desire Beyond Belief. *Australasian Journal of Philosophy*, 82(1), 77-92.
- Lewis, D. (1976). Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review*, 85(3), 297-315.
- Lewis, D. (1986). Probabilities of Conditionals and Conditional Probabilities II. *Philosophical Review*, 95(4), 581-589.
- Lewis, D. (1988). Desire as Belief. *Mind*, 97(387), 323-322.
- Lewis, D. (1996). Desire as Belief II. *Mind*, 105(418), 303-313.
- Milne, P. (1997). Bruno de Finetti and the Logic of Conditional Events. *The British Journal for the Philosophy of Science*, 48(2), 195-232.
- Oddie, G. (1994). Harmony, Purity, Truth. *Mind*, 103(412), 451-472.
- Oddie, G. (2001). Hume, the BAD Paradox, and Value Realism. *Philo*, 4(2), 109-122.
- Price, H. (1989). Defending Desire-as-Belief. *Mind*, 98(389), 119-127.
- Stefánsson, H. O. (2014). Desires, Beliefs and Conditional Desirability. *Synthese*, 191(16), 4019-4035.
- Weintraub, R. (2007). Desire as Belief, Lewis Notwithstanding. *Analysis*, 67(2), 116-122.

Following Desire as the Ethical Postulate of Psychoanalysis

Bogna Choińska

Pomeranian Academy in Slupsk

Abstract. One of the most controversial theses of Jacques Lacan is his conviction that a very specific relationship links ethics and desire. The aim of the article is to present what this new relationship consists in, and, further on, to outline the weaknesses of this concept, which does not take into account the existence of the sovereign good as a category available to cognition. According to my thesis, Lacan believes that the ethics of Supreme Good, or simply traditional ethics of goods, leads the human subject to remain, voluntarily (and perhaps thoughtlessly), within the Imaginary dimension. The idea of the ethical postulate will be treated here not so much as something applied during psychoanalysis, but as a general clue as to how people should behave.

Preliminary remarks

One of the most controversial theses of Jacques Lacan is his conviction that a very specific relationship links ethics and desire. The aim of the article is to present what this new relationship consists in, and, further on, to outline the weaknesses of this concept, which does not take into account the existence of the sovereign good as a category available to cognition. According to my thesis, Lacan believes that the ethics of Supreme Good, or simply traditional ethics of goods, leads the human subject to remain, voluntarily (and perhaps thoughtlessly), within the Imaginary dimension. The idea of the ethical postulate will be treated here not so much as something applied during psychoanalysis, but as a general clue as to how people should behave. It also seems reasonable to draw attention to the understanding of the word „Other,” which I will use here. This concept can be understood in different ways, as the history of philosophy shows. For example, Emmanuel Levinas calls the „face” of another person the Other, and in his thought the relationship with the Other consists in the immediacy of experience, where the Other is the same as man or God (of course, I present it in very simple terms, only for the purpose of contrasting it with Lacan's attitude). For the French psychoanalyst, the Other refers to the symbolic order — it is language and a set of social rules. Lacan wrote: “Before human relationships are strictly established, some relationships are already determined” (Lacan, 1973, p. 18). Therefore, the Other is not

another subject (God, man, animal), but an order of mediation not existing in reality. Human relationships do not exist other than mediated by the Other, and, therefore, they are always complex and full of confusion (hence the thesis about the non-existence of sexual relation¹). Such an understanding of the Other involves the issue of the triadic division of human “willings” into: need (*besoin*), demand (*demande*), and desire (*désir*).

Desire and demand

Lacan argues that ethical action consists in “realizing one's desire,”² which means, at the same time, acting according to specifically understood Symbolic law. Lacan clearly emphasized: “This is so evidently the case that one can, in short, say that the genesis of the moral dimension in Freud's theoretical elaboration is located nowhere else than in desire itself” (Lacan, 1997, p. 5). How to understand this enigmatic statement? First, let us focus on the difference between desire and demand. According to Lacan, a human being lives in three dimensions, which he calls the Imaginary, the Symbolic, and the Real.³ The Imaginary might be determined by knowledge (about oneself and the world), the Symbolic — by the Other, and the Real — by what exists like the Kantian noumenon, and this noumenon is not only cognizable but also traumatic. The three orders “function” in a Borromean knot, which is so intertwined that the disruption of one of the loops releases the others. This means that man cannot be described by a single order. Man is not a creature who will remain only in the Imaginary or only in the Symbolic or only in the Real. Lacan pointed out at one of his seminars: “There is no way one can reduce desire in order to make it emerge, emanate, from the dimension of need” (Lacan, 1997, p. 169) because need is the condition of a demand. Demand comes from the Imaginary and has a particular place in the subject, associated with this particular order. Demand is specific: when I am hungry, I demand good food. The question arises: why do I demand food instead of just eating? Demand differs from need by the fact of approaching the Other (that is, in this context: „I am hungry — feed me”). Demand cannot be reduced to need, and this is due to the fact that the subject's being continually intersects with the field of the Other. Demand is defined by the fact that there is the Other. When my demand is satisfied, I feel a momentary sense of satisfaction and I often feel that it is too small anyway. The food could still be better, better served, etc. Therefore, despite satisfying a demand, and thus meeting a specific need, the satisfaction is still not enough. This

¹Lacan (1999a, p 47).

²This thesis often appears in Lacan (1997).

³This is completely explained by Lacan in Lacan (2015)

constant hysterical questioning, this constant „that is not this” (the subject as such is a hysterical subject for Lacan). In other words, satisfying a demand may take place but only in the register of the Imaginary. At this point, the matter of subject's ethics would not have any foundation at all, i.e., if the subject only existed in the order of the Imaginary. Demand as such does not involve ethics in the subject. Paradoxically, however, desire is related to demand, because it is defined as the difference between need and demand — a desire is a negativised demand. It happens because the desire is only expressed by the demand, so to speak: the demand is not direct. The desire somehow „degrades” itself to the demand, but the content of demand does not constitute the content of desire. This sometimes happens with the demands of a psychotic subject (and more specifically — a psychotic subject does not actually have desires, only demands), but that's not the subject of my considerations here. Apart from exceptional cases, there is no situation in which a human being only expects to satisfy demands. Satisfying a demand is always accompanied by something more, and this „more” is given or not, as if involuntarily. I am referring again to the example of the demand for a meal. The child says that he/she is hungry. However, quite often, the reason for the demand is not the actual need, but a desire for love and intimacy with the mother. The meal served by the mother (or sometimes not served) stimulates the child to interpret the desires of the mother. Satisfying the demand (or not), the mother can offer dinner to the child as: evidence of concern, even a kind of punishment or reward, or a proof of love. At the same time this concern, punishment, or reward is a matter of interpretation of the mother's desire by the child. The child's desire introduces this child into the relationship's circle of the mother's desires. Therefore, the served meal does not only satisfy the need evoked by hunger and the demand for food, but it starts an endless game, the dialectics of desire and demand. In this sense, desire is always the desire of the Other and cannot be of different character. Desire is this part which is formed after the demand passes through the field of the Other.

Desire as desire for desire and defense against *jouissance*

Lacan defines desire as “a desire for unfulfilled desire” because it is related to the Symbolic. We can hear the echo of the Hegelian dialectics of Master-Slave, but, for Lacan, synthesis is not achieved.⁴ Desire cannot be fulfilled due to the fact that this is desire of the Other in every possible sense: I desire the desire of the Other, I desire the Other to desire me, I desire

⁴Lacan was (together with many other French intellectuals) a dedicated listener of Alexander Kojève's lectures on Hegel. These lectures are available in English: Kojève (1969).

the Other by myself, I desire to be in the desire of the Other, I desire the Other to be in my desire, etc. When I desire something, it is not that I want to satisfy my need (which may be, in the common opinion, either good or bad), but I am trying to guess what the Other wants, how to become desired by the Other; his/her desire is a mystery for the subject, which makes insisting on the desire an attempt to capture what the Other wants from me. For further clarification: Lacanian “Que vuoi?” means “what do you want?” rather than “what do you want from me?.” If the subject expresses its demand towards the Other, it behaves neurotically: a neurotic person wants the Other to tell him/her what he wants from the subject; the neurotic does not want to solve the mystery of the Other's desire, he/she wants the Other to give straight answers. But such an answer does not exist (in terms of the Hegelian synthesis).

Still another definition of desire is describing it as refraining from satisfaction. Lacan calls the desire a defense against *jouissance*, which is instinctive, and, as such, beyond good and evil. Now, I will try to explain what Lacan means by *jouissance* and how it relates to the question of desire. *Jouissance* is the satisfaction of a drive that accompanies every activity of the subject, and this drive has only got one aim: to „take delight.” Well, Lacan in his famous text of *Écrits II* ⁵ postulates the “equality” of all acts of *jouissance*. It means that both the subject doing evil (e.g. killing animals for pleasure) and the subject doing good (in this context, it could be a defender of animals' rights) will experience the same satisfaction from the act of *jouissance*. Why are these acts morally indistinguishable? This is proved by the fact that those who support the good which is widely accepted by the society will automatically exclude those who support the opposite (they do evil, they are torturers, murderers — for example of human fetuses, animals, etc.). A common result of such exclusion is opposing the “sinners” by taking the place of the previous “sinners.” In this way, as Lacan writes, stakes are burning where the bad are sacrificed for the so-called common good, but they are sacrificed using the methods of hated executioners. This shows how variable *jouissance* is and how easy it is to go from doing good to doing evil. That is why, as I wrote, it is beyond ethics, beyond good and evil. And therefore every type of *jouissance* (even an animal rights activist who devotes his/her money and time to this activity) is dangerous, as it can quickly reveal its “evil” instinctual face to the subject.⁶ Desire is not associated with evil, because doing evil is connected with a specific deed performed at a definite time

⁵See Lacan (1999b, pp. 243-273).

⁶Peter Singer wrote in this context about the acts of violence connected with defending animal rights. Obviously, these acts were condemned by Singer, using arguments which coincide with mine as presented in this article, see Singer (2004, pp. 24-25).

(a deed which is a good for the subject, provided that we leave behind the naive understanding of the good as something positive in the opinions of civilized societies) and can only respond to the demand of the subject. A deed dictated by the knowledge of the good looks like a satisfied demand — it is complete and finished, and it gives some satisfaction. But satisfaction is addictive because there is never enough of it — there is a need, therefore, to perform new deeds in the name of the “good,” the need “of one more effort” as in Marquis de Sade.⁷ The infinity of *jouissance* is not the same as the infinity of desire for two reasons. First, simply because it is *jouissance*, and so something that belongs to the instinctual and non-ethical sphere. Second, because it does not attempt to deal with the structural deficiency of the human being (more on that later) and does not try to complete it, as „lighting stakes” in the name of a deliberately established good may increase injustice, sometimes more than the omission of such actions. Meanwhile, desire (another definition) must be based on the knowledge of its *jouissance*; this is an absolutely necessary condition. To get involved in the game of desire, my desire and Others’ desire, is to enter an unending game. The ordinary hurting of a living creature, whether for material benefit or through the discharge of aggression, is, similarly to eating a meal which only satisfies hunger, the completion of a demand, the same as planting a bomb at an abortion clinic (in good faith, of course). Desire is not satisfied with a single instance of fulfillment, because that fulfillment leads one closer to the inability to satisfy desire. Desire is the desire for the Other, and it is therefore the desire for an unfulfilled desire. We will never know what we want or what the Other really wants from us or what we can offer him as Symbolic beings. Desire is almost Pythian, because we cannot satisfy it completely when we do not know what exactly it concerns. Therefore, evil, as an egoistic desire for still more objects, is not, according to Lacan, located in the Symbolic dimension, but merely in the Imaginary, and thus it is always related to demand and never to desire. Also, only at the level of the Imaginary does a particular object appear as a reason (in the Symbolic register, it is an effect) for a demand. At this level, the principle of pleasure is in force, in opposition to morality and to the order of the Real. Demand is demanding of pleasure and is accompanied by the knowledge of what this pleasure consists in, what object it needs, and how to achieve it. However, if we agree with Lacan that “knowledge is the appearance of truth,” (Lacan, 1999a, p. 135) it turns out that we only have false knowledge about the goods that we want to have and give. But concerning desire, we cannot even obtain such apparent knowledge. However mysterious it may sound, desire has to be “retrieved”: as if “pulled out” of the unconscious with

⁷ Sade (1970), pp. 81-112.

the help of words. Therefore, postulates or rather guidelines for the direction of desire penetrate one's consciousness with difficulty, and this process is connected with moral development of the individual. In this sense, morality has to do with the truth of the subject — the truth about the unconscious of the drive, about its own *jouissance*.

Desire as filling in the lack

What is the role of desire when it comes to *jouissance*? One must recall here the Lacanian teaching about structural lack in the Symbolic order. The lack (referred to as *phallus*) makes the world always appear defective. Its root cause, however, does not lie in the nature of the world, but in the incompatibility between humans and the world. Man is the only creature who lives not so much in the world, but in language. And because, firstly, language creates the ontology of our being, and, secondly, language has a hole⁸ which is impossible to fill in, to us the world will always seem unjust, merciless, inhuman, etc. All that the attempts to repair the whole world do is allow the repairers of all sorts to experience pleasure, and they lead to further manifestations of evil. Real good, according to Lacan and as I have already mentioned, is not an answer to the need produced by demand. No object of need is the good. It cannot be so, because it is a necessity. There is also no knowledge of the needs of beings other than the subject himself. In contrast, the objects of desire are only a possibility, they appear metonymically. Desire is always metonymic.⁹ No object will meet the expectations of man's desire, because it is never the right object. Objects are apparent. Desire is realized through the Symbolic. This means that the aim of desire is not an object, but that the latter (according to Lacan's terminology, 'objet petit a') is produced as a kind of surplus in the relation of the subject to its desire (Lacan defines 'phantasm', that is, the function of the subject in relation to its desire, as: $S \diamond a$, where the crossed out "S" signifies the position of the subject towards the object, " $\diamond$ ", is the relation of desire, and "a" is the object). In other words: the object is a result of desire, not the reason for it, or more precisely: desire „has" its object which is at the same time its cause, while there is no object

⁸ The hole that we find in language is a result of the otherness of language vis-a-vis the world. The language and the world are two different entities, but a human being „lives" only in language, and therefore his/her direct contact with the world is impossible. I expand on this in Choińska (2014), chapter III.

⁹ Lacan describes metonymy with the help of the following equation: $f(S \dots S')S = S(-)s$, which means that the line between the signifying and the signified is not crossed. And metaphor is presented in other equation: $f(S'/S)S = S(+)s$, which means that the line is crossed, which enables creation of new meanings which are called creating or poetical by Lacan. See Dor (1998, pp. 52-53).

that would be the aim of desire — desire aims not for an object, but for satisfaction; that is why objects are not only apparent, but also variable. Desire, in contrast to demand, is an attempt to patch up the Symbolic hole, and therefore it is an infinite process. It is unethical not to desire, not to insist on desire, because when we cease to desire, the *jouissance* comes onto the stage and reigns.

Object *a* is only an appearance, and hence desire cannot be a wish to do evil. ‘Petit *a*’ on the Imaginary level is associated with the fear of aggression by some presumed other “I,” wanting to rob us of our good.¹⁰ Also, it is just a picture of the good belonging to me as an ideal (narcissistic) “I,” so in fact, just a mirage of “I.” Lacan talks elsewhere¹¹ about the reduction of ethical order to social coercion. This can be seen particularly clearly, in my opinion, when social coercion of late capitalism takes on the form of generating demands and encouraging endless purchases of certain goods. On the other hand, in order to sustain desire it is necessary to leave behind goods understood as Imaginary objects, hence the connection between desire and ethics, one of the features of which, as Kant already pointed out, is disinterestedness. In this sense, Jonathan Safran Foer wrote about the human necessity to deny oneself a (e.g. non-kosher) meal, even during extreme famine.¹² He wrote about his grandmother, who was starving at the end of the war, but when a Ukrainian peasant brought her a piece of pork meat, she refused to eat it, explaining: “If nothing matters, there is no point in saving anything” (Foer, 2013, p. 20).

Lacan's Ethics. Ethics of desire for death?

Desire, through its emphasis on satisfaction, complicates traditional ethics, and Lacan's ethics is sketched out somewhat across traditional ethics. This is because the latter has a problem with calling the subject's satisfaction (that is: my own) a good. If satisfaction is taken into account at all in classical ethics, it is only as the satisfaction of the Other. And Lacan claims that the satisfaction of the Other is merely our interpretation of it. We can only feel our own satisfaction. And so our own satisfaction should be what we deal with, e.g. through neutralization, limitation, etc.

In this context, it becomes necessary to clarify the relationship between desire and law, and it can be bluntly said and that desire is simply the law. Moral law is articulated in relation to the Real (Thing, *das Ding*). Sade's law is the opposite of the Kantian imperative: this is ethics without feelings (of mercy or compassion). Sade's imperative is: “Take delight” regardless of whether it makes you feel pleased or not. The Kantian imperative

¹⁰ Lacan raises this topic in *Le stade du miroir*, in Lacan (1966).

¹¹ Lacan (1997, p. 240).

¹² Foer (2013).

is: „Follow your duty,” regardless of the fact that emotions often dictate the violation of this law. Therefore, both Sade and Kant, in the interpretation of Lacan, try to persuade us to take instinctual delight, although they seemingly stand at opposite poles. For, would the punishment for non-compliance with the Kantian duty be not performed with equal delight as the obligation itself? According to the French psychoanalyst, in both of these conceptions there is not enough knowledge about the „horror” of *jouissance*. Meanwhile, desire is always a desire for death — it is a consequence of his theory of desire that it cannot stop even for a moment. There are no unethical desires. Eventually, desire (as I understand it) is fulfilled in the moment of complete dedication. Therefore, the one who continues to desire is a Hegelian Master — and is not afraid of death. Desire as desire for death is also related to the fact that there is no mediation between the order of the Real (Thing) and pleasure. What comes from *jouissance* is a prohibition and transgression of *jouissance* itself, and this transgression is only possible if it is supported by the Symbolic law.

If we follow a deconstructive understanding of Fromm's opposition of being and having,¹³ we will admit that being might also be a kind of possession, and then, in this context, we can understand the attitude of Antigone (I will return to the problem of Antigone later). According to Lacan, “Antigone appears as *airrovogos*, as a pure and simple relationship of the human being to that of which he miraculously happens to be the bearer, namely, the signifying cut that confers on him the indomitable power of being what he is in the face of everything that may oppose him” (Lacan, 1997, p. 224). If we assume that life is a kind of good we own, then the idea of resignation from life in favor of sustaining the desire to be something more than an owner explains partially the attitude of Antigone and her readiness to die. Lacan wrote something similar about the Cathars:

they had basically to desist from any act that might in any way favor the perpetuation of the world, considered as execrable and bad in its essence. The practice of perfection thus consisted essentially in seeking to achieve death in the most advanced state of detachment, which was a sign of reintegration into an Edenic world characterized by purity and light, the true world of the original good Creator, whose creation had been sullied by the intervention of the bad Creator or Demiurge (Lacan, 1997, p. 103).

To forestall the interpretations (as to the intentions of Lacan as a singer of death) evoked by Polish philosophers,¹⁴ I must add an important issue here. Lacan clearly distinguishes ethical actions associated with desire and leading toward the triumph of death from those which he called,

¹³ See Choińska (2012).

¹⁴ See Bielik Robson (2008) and Dybel (2009).

probably inspired a bit by the language of Heidegger, a triumph of Being toward death. There are two ways of discussing (in one of the seminars) the problem of Antigone, who, in choosing death and relying on her own desire, gives up the goods which would keep her from realizing her desire. In contrast, the triumph of Being toward death is simply the triumph of life (Lacan writes here about those undergoing psychoanalysis who managed to go through it successfully and change their lives, by no means in the direction of allowing death to triumph).

While fulfilling one's demands (in certain contexts we could even call them whims, more or less dangerous to others) means satisfying one's needs and is associated with the principle of pleasure, paradoxically, the infinite following of desire touches the Real and the satisfaction of one's drives. That satisfaction which Freud described as a satisfaction so unwanted in our dreams that we wake up in horror.¹⁵ Fulfilling a desire encounters two obstacles: the first is the horror of *jouissance* that the desire „is spinning” around, and which we can determine at the level of the subject as the threat of pain and the fear of the unknown. The second obstacle is the beauty filling up our phantasm, which disarms our desire. This is a situation in which we are afraid to „touch” our desire, and what appears in place of a confrontation with it is art, poetry, or, when it's an average analysant — not an artist — references to them.—If the way of an individual to becoming ethical does not consist in the fulfillment of demands, and the way of desire must be “retrieved,” beauty can somehow make this last thing impossible, because it allows the subject to experience the feeling of the sublime (other than the sense of the good). Nevertheless, both the desire of an uneasy satisfaction related to *jouissance*, which is the mystery of the Real, and the sense of beauty or sublimity, are equally an obstacle to accessing the subject's desire and a kind of bait, a point for the enigmatic desire to take hold. Pain opens us to the limit of the Thing before its absolute neither good nor evil. As for *jouissance*, there is a phenomenon that Lacan calls the “tyranny of memory,” where memory is a rival of ordinary pleasure and also satisfaction. This is the memory of the Real, of trauma, of life and death. The subject is (to paraphrase here the famous phrase by Bielik-Robson) “an animal that forgets too easily,” but must become responsible also for what it has forgotten. Beauty, on the other hand, may turn out to be “the splendor of truth”; it sublimates our drive and satisfies it somehow. “Sublimation is the satisfaction of the drive with a change of object, that is, without repression” (Lacan, 1997, p. 230). The situation of the good as satisfaction and of beauty as a phantasm, both of which limit and, at the same time, enable access to desire, is, as we can see, paradoxical.

¹⁵ Freud (1996).

Lacan (following Freud) distinguishes between primary processes, dominated by the identity rule of pleasure (there is only pain and pleasure here, but pain is to be avoided) and also by the knowledge of the goods which the subject wants to have and bestow onto others, and secondary processes, determined by the rule of reality, where perception lies (rather than thinking). Only at this level does the key to the ethics of psychoanalysis appear. The secondary processes, which we access through certain known words, allow us to combine these with the unconscious in such a way that they can move into the chain of meaningful associations appearing already at the conscious level. In this sense, desire cannot be based on knowledge — on the contrary, it should be based on: first, the feeling of one's own satisfaction while, second, this satisfaction cannot be combined with an ordinary surrender to *jouissance*. Desire protects against pleasure connected with the drive thanks to the intention of the one trying to "patch up" the lack in the world, because from the point of view of the ethical subject there is too little good in the world, too little justice, etc. Thus, desire operates in the sphere of the Symbolic.

The drive has two dimensions. It tends to reach entropy balance in non-organic matter. In the dimension of history, the drive is in the chain of the signifying, and its manifestation begins with the first word, which, touching other words, will create the "personality" ("extimacy") of the subject, that subject taking its own road toward death. Antigone defying Creon in the name of blood ties with her dead brother is, as Lacan says, the image of barbaric mercy. Antigone's desire is destructive; it can be contrasted with the wicked lust of her mother — Jocasta. Antigone is the guardian of a dead criminal on account of the memory of individual signifiers — subjects: brothers, one of whom could experience the so-called „second death,” being crossed out of the chain of the signifying. That is why Antigone's attitude is a revelation, shining with splendor. Antigone cannot agree to Creon's recommendations and become the one who serves the goods (the common good of the society). Then she would have committed the sin of forgetting about the bonds which had located her and her family within the chain of the signifying. Beyond this chain — of Antigone as a heroic creature — she does not exist as a subject. This is very important: breaking the Borromean knot and crossing the subject out of the Symbolic reduces him or her to the Real, that is, to the Thing. And that makes this subject disappear, reducing it to a „whirl” of drives, to what is beyond good and evil.

If sublimation involves only the change of the object, then Antigone changes the object in itself, not exchanging it for another, but, as Lacan claims, from the point of view of the "Last Judgment," no object can be in this case better than another. Thus something completely new must

be involved here — a change of attitude. Antigone is ready to die. The lesson psychoanalysis teaches us is that the idea of a sovereign good is based on demands, on the order of the Imaginary, and therefore it shows their futility. Creon is the one who demonstrates the impossibility of, or rather the leaving behind of, this desire. That is why his defeat is of a different kind: it is the sheer triumph of death. Creon is entangled in traditional ethics; he is concerned with serving the goods, which, in his case, is strongly connected to the role of power he represents. But authority is not interested in the desires of subjects (“it always makes desires wait”), which is particularly visible in the authorities in the capitalist system (or in the authority of capital, or of capitalistic discourse): all it can offer to its citizens is showering them with various goods, as if from the horn of plenty. Citizens are persuaded to buy these goods with the help of advertising, “scientific” opinions, etc. In conclusion: Lacan opposes serving the goods to following one’s desire.

Psychoanalysis attempts to restore to the subject its relationship with desire; analytical activity is strictly ethical and in the service of ethics; healing happens as if incidentally, it is a result of ethics, and not of demands. The subject has to position him/herself in the whole tragic sense of life, rather than merely choose the goods (for him/herself and others), because they only bring a momentary joy. They do not give the subject the ability to self-actualize, making it impossible to “cure” either neurosis, or perversion, or even so-called existential problems. Thus, the greatest fault is to stop „insisting on one’s desire,” i.e., betraying one’s own desire. When we give up our desire, it is only in the name of some goods. But whose goods are they? How do we know what others expect, since, even when they tell us what they want, or (as in the case of animals) when we guess what they want, we only touch the “appearance of truth”? Desire has something of the Real — it stubbornly places us again and again on the same path towards its fulfillment. And if desire is not finite (in contrast to satisfying hunger), we can say that it never gives up and thus it constitutes a triumph of Being toward death (with special emphasis on Being). With desire, we pay off the debt in the form of mercy or compassion. And when we give up desire, we betray ourselves and others.

Desire is a “metonymy of our being”; the good can only serve as payment for satisfying a desire. In fact, desire can even overcome fear and pity, not by following the good (because it does not know anything about it), but by creating it as it fills the lack in the Symbolic. However, knowledge or memory about the impulsive nature of human beings is necessary here. Knowledge and memory not about the sovereign good, but about the fact that the subject is also a Thing out of the order of the Real.

Final remarks. Introduction to a critique of Lacan's ethics

The departure from the concept of the sovereign good, and particularly from the theory of values which exist in a transcendent way, raises serious doubts. First of all, the area of ethics, deprived of absolute (or even any) knowledge of the good — and this attitude is not only typical for Lacan, but for the whole poststructuralist formation (Derrida, Foucault, Lyotard)¹⁶ — remains a “wild” area. In my opinion, such an understanding of ethics is not enough to guarantee ethical behavior, as what is assumed here is the possibility of completely distancing oneself from knowledge about the good and relying instead on acts based on the satisfaction of desire. Does this exclude unworthy behavior, especially when we take into consideration not only the development of a specific subject, but the idea of moral progress of the whole society? I do not think so. For this larger area, in my opinion, cannot remain deprived of a hierarchy of good and evil based on the knowledge that every subject acquires while growing up in a society. This means, in turn, that such knowledge also functions in psychoanalysis, but in a somehow unconscious and “thoughtless” way. Therefore, even if we have high self-awareness (in the psychoanalytical sense), we still have to make ethical life decisions, endlessly. After all, abandoning knowledge about both the object of the good we are aiming at and the reasons for our decision creates a situation where we have to rely only on a „wild” act dictated by... we do not actually know what. In my opinion, such an act, despite assurances from the followers of poststructuralist ethics about its “purity,” remains entangled in the theory of values, even though it is consciously rejected. And if this theory is rejected, it is not clear whether it does not get replaced by superstitions and prejudices that may cause unconscious evil. As it is not possible to discuss all my doubts within the scope of this article, I can only signal these difficulties here. For the purpose here was to present Lacan’s position, as his theory deeply influenced thinking about ethics within the whole post-structuralist tradition. My critique will be developed in subsequent publications.

¹⁶ See my work Choińska (2014).

References

- Bielik-Robson, A. (2008). *Na pustyni. Kryptoteologie późnej nowoczesności*. Kraków: Universitas.
- Choińska, B. (2012). Poststrukturalizm a myśl etyczna Fryderyka Nietzschego. *Filo-Sofja*, 16, 183-194.
- Choińska, B. (2014). *Podmiot i dyskurs w świetle myśli wybranych przedstawicieli poststrukturalizmu francuskiego*. Kraków: Universitas.
- Dor, J. (1998). *Introduction to the Reading of Lacan. The Unconscious Structured Like a Language* (S. Fairfield, Trans.). New York: Other Press.
- Dybel, P. (2009). *Okruchy psychoanalizy. Teoria Freuda między hermeneutyką a poststrukturalizmem*. Kraków: Universitas.
- Freud, S. (1996). *Objaśnianie marzeń sennych. Dzieła*, vol. I (R. Reszke, Trans.) Warszawa: KR (English translation: Freud, S. (2010). *Interpretation of Dreams* (J. Strachey, Trans.). New York: Basic Books).
- Foer, J. S. (2013). *Zjadanie zwierząt* (D. Dymińska, Trans.). Warszawa: Krytyka Polityczna (Foer, J. S. (2009). *Eating Animals*. New York: Little, Brown and Company).
- Kojeve, A. (1969). *Introduction to the Reading of Hegel: Lectures on the Phenomenology of Spirit* (J. H. Nichols, Jr., trans.). Ithaca: Cornell University Press).
- Lacan, J. (1966). *Ecrits*. Paris: Le Seuil.
- Lacan, J. (1973). *Les quatre concepts fondamentaux de la psychanalyse* (S XI). Paris: Le Seuil.
- Lacan, J. (1997). *The Ethics of Psychoanalysis* (SVII), transl. D. Porter, London, New York: W.W. Norton and Company.
- Lacan, J. (1999a). *Encore* (S XX) (limited internet version), following English edition, J. Lacan, *On Feminine Sexuality. The Limits of Love and Knowledge*.
- Lacan, J. (1999b). Kant avec Sade. In *Ecrits II*.
- Lacan, J. (2015). *RSI* (S XXII) (typescript made accessible by NLS).
- Sade, D. A. F. (1970). Francuzi, jeszcze jeden wysilek, jeżeli chcecie stać się republikanami (J. Lisowski, Trans.). *Twórczość*, 2, 81-112. (English translation: Sade, D. A. F. (1966). Yet Another Effort, Frenchmen, If You Would Become Republicans. In Sade, D. A. F. *Philosophy in the Bedroom and Other Writings* (pp. 296-384) (R. Seaver & A. Wainhouse, Trans.). New York: Grove Press).
- Singer, P. (2004). *Wyzwolenie zwierząt* (A. Alichniewicz, A. Szczęsna, Trans.). Warszawa: PIW (Singer, P. (1975). *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York: HarperCollins).

Grounding Reichenbach's Pragmatic Vindication of Induction

Michael J. Shaffer

*Department of Philosophy
St. Cloud State University*

Abstract. This paper has three interdependent aims. The first is to make Reichenbach's views on induction and probabilities clearer, especially as they pertain to his pragmatic justification of induction. The second aim is to show how his view of pragmatic justification arises out of his commitment to extensional empiricism and moots the possibility of a non-pragmatic justification of induction. Finally, and most importantly, a formal decision-theoretic account of Reichenbach's pragmatic justification is offered in terms both of the minimax principle and of the dominance principle.

1. Introduction

Hans Reichenbach's pragmatic treatment of the problem of induction (presented and developed in his 1949a, 1932/1949b and 1949c) was, and still is, of great interest. However, various influential commentators have dismissed it as a pseudo-solution and/or regarded it as problematically obscure.¹ This is, in large part, due to the difficulty in understanding exactly what Reichenbach's solution is supposed to amount to, especially as it appears to offer no response to the inductive skeptic. As Laurence Bonjour claims,

...the significance of Reichenbach's pragmatic justification remains obscure. As he himself insists, that justification still yields no reason at all for thinking that inductive conclusions, or any of the myriad further beliefs which are epistemically dependent on them, are to any degree likely to be true. The sort of justification in question is thus not epistemic justification, as that concept was construed above; to show that beliefs are justified in this alternative way does not answer, or even purport to answer, the basic skeptical worry about induction, and is indeed quite compatible with the deepest degree of skepticism. It is thus hard to see why it should be regarded as any sort of solution to the classical problem of induction (Bonjour 1986, p. 99).

¹ See, e.g., Skyrms (1966), Salmon (1966), Bonjour (1986), Bonjour (1992), Bonjour (1998), Rosenkrantz (1981), and Kelly (1991).

This sort of dismissal of Reichenbach's views on induction surely depends to some degree on his adherence to the controversial frequency interpretation of the concept of probability.² This is plausible to assume in light of the wide-spread popularity of subjectivism about probability that has dominated probability theory subsequent to Reichenbach's work and this aspect of Reichenbach's approach to induction has been treated at length in a variety of other places.³ Another aspect of his vindication that likely fuels the charge of obscurity is the failure of these critics to pay more attention to Reichenbach's commitment to a purely extensional metaphysics. Finally, Reichenbach's attempt to rationally ground induction on the basis of purely pragmatic considerations is also likely to have, at least in part, given rise to the charges of obscurity and failure.⁴ This last point is especially important because Reichenbach is rather cavalier and informal in his assertion of the claim that following the inductive rule is the best thing to do from a pragmatic perspective.

First, it will be shown here that what Reichenbach does in his later work on induction is to establish an important epistemic limitation of extensional empiricism and that there really is nothing especially obscure about Reichenbach's thoughts on induction in this respect. He was simply working out the limits of extensional empiricism with respect to inductive inference. In fact, Reichenbach conveys this very point to Russell in an important letter in 1949.⁵ It will be shown that these aspects of Reichenbach's position lead directly to his assertion that the only manner in which the inductive rule can be justified in the primitive state of knowledge prior to making sequences of inductive inferences is pragmatic in nature. Finally, and more constructively, it will be shown here that there are perfectly coherent, formal and pragmatic justifications of Reichenbach's inductive rule in the primitive state of knowledge in the forms of the maximin rule for decision making and the dominance principle. With this account of the pragmatic justification of induction in hand, we will see

² Hájek (1997, 2009).

³ See, e.g., Skyrms (1966), Salmon (1963), Salmon (1966), Galvaotti (2011) and Teng and Kyburg (2001).

⁴ The sense of pragmatics used here is just the idea that there are justifications that are non-epistemic (i.e. not related to truth, approximate truth or probability) and which are based on some more or less well-understood notion of instrumental success or utility.

⁵ The letter is a response to Russell's criticisms of Reichenbach's approach to the problem of induction as presented in Russell's *Our Knowledge of the External World*. This approach stands in sharp contrast to Reichenbach's early work which resembles Russell's approach in holding that the principle that grounds induction is synthetic a priori. See Eberhardt (2011) for discussion of Reichenbach's early views. He explicitly rejects this view in his later work and tells Russell in a 1949 letter that "Induction does not require an intensional logic" (1949d, p. 410).

that Reichenbach's justification of induction can be given a principled ground.

1.1 Overview

Let us begin by recalling that Reichenbach's pragmatic justification of induction is based on the following (reconstructed) line of argumentation (i.e. *the basic Reichenbach argument*):

P1: Either nature is uniform or it is not.

P2: If nature is uniform, then scientific induction will be successful.

P3: If nature is not uniform, then no method will be successful.

∴ If any method of induction will be successful, then scientific induction will be successful.⁶

But, according to Reichenbach and echoing Hume, we cannot know whether nature is uniform or not, because it is neither a matter that can be settled *a priori* nor is it a matter that we can non-circularly establish *a posteriori*. So, as Reichenbach sees it, although we know that if any method is successful, then scientific induction will be successful, we cannot know that any method really is successful. The gist of his attempt to justify inductive practice then comes from the idea that while we do not know that any method will actually be successful, we also do not know that no method will be successful. Given this result and the fact that scientific induction can be shown to be an optimal method (in this important sense of "optimality"), we ought to accept induction as being justified, at least pragmatically speaking. As we shall see, what is at the heart of this view is Reichenbach's metaphysical commitment to a form of extensional empiricism that tolerates only the existence of particulars.

In any case, as Salmon correctly pointed out in his 1966, the Reichenbach argument depends on a false dichotomy. The uniformity of nature is, of course, not an all or nothing matter. We can, of course imagine possible worlds that contain only individuals with degrees of uniformity that vary radically. So, the uniformity of nature seems to be a matter of degree, and it is at least plausible to believe that there is a measure of the uniformity of extensional worlds. If this turns out to be viable, given the space of possible worlds U , we could define a measure $m(x)$ on U such that $m(x)$ maps the elements of U into the continuous open interval $[0,1]$

⁶ This presentation of a simplified version of Reichenbach's main argument is taken from Skyrms (1966). It is important to note at this juncture that the various criticisms of Reichenbach's views, other than Russell's, will (for the most part) be ignored here. To address all of those criticisms would require too much space, and the point of this paper is more historical in any case.

representing the uniformity of that extensional world. This suggests that Reichenbach's attempt to justify induction needs to be retooled in order to accommodate a concept of world-uniformity that admits of continuous degrees. When this is done we can usefully reformulate the basic Reichenbach argument as follows. Consider our world w_a (the actual world), where $w_a \in U$, with a fixed but unknown measure of uniformity, the set of all inductive methods $\mathcal{Y}$,⁷ where $y_i \in \mathcal{Y}$ and such that each inductive method has a probability of arriving at a true conclusion in its domain of application, a function $f(m(w_n), y_n)$ that maps worlds with degrees of world-uniformity and inductive methods into the space of probabilities,⁸ and a constant λ that represents the chance probability of an inductive method succeeding at a world.⁹ If we understand ε as the degree of world-uniformity required for any inductive rule to be reliable with a reliability greater than chance, i.e. greater than λ , then the more *sophisticated Reichenbach argument* can be stated as follows:

P1': If the probability that $m(w_a) = 1$ is 1, then scientific induction will be successful.

P2': If it is probable that $1 > m(w_a) > \varepsilon$ with probability less than 1 but greater than λ , then scientific induction will be successful with probability p , where $p > \lambda < 1$.

P3': If it is probable that $\varepsilon > m(w_a) > 0$ with probability greater than 0 but less than λ , then scientific induction will be successful with probability p , where $p < \lambda < 0$.

P4': If the probability that $m(w_a) = 0$ is 1, then no inductive method will be successful.

$\therefore$ If any inductive method will be successful, scientific induction will be successful.

It should be noted that Reichenbach's conclusion still holds in this case, and we will consider the significance of this conclusion in what follows.

⁷ Inductive methods are, simply, rules for accepting conclusions concerning unobserved cases based on observed cases.

⁸ The function $f(m(w_n), y_n)$ seems, intuitively, to be a natural sort of function, as degrees of world-uniformity seem to be closely related to the probability with which a method produces true conclusions. What $f(m(w_n), y_n)$ is supposed to yield is a probabilistic measure of the general reliability of a given method at a world with a given measure of uniformity, and, as we shall see in section 4, what this function really represents is the set of worlds where an inductive method with a well-defined probability of arriving at the correct value of a stable frequency will actually produce the correct values.

⁹ In other words, λ represents the threshold at which methods are no better at producing true conclusions than randomly selecting conclusions from the set of all statements of a given language $\mathcal{L}$, and, as we shall see, a method that performs at a success rate no better than chance is no method at all. However, the general successfulness of an inductive method will turn out to be a more complex matter involving two aspects. The first concerns the reliability of the procedure in its domain, and the second concerns whether there exist elements of that domain at a world.

However, before we proceed to do so, it will be instructive to reconstruct Reichenbach's treatment of induction in much greater detail in order to see just what it amounts to and what it implies about inductive inference.

2. Reichenbach's Conception of Scientific Induction

The primary motivation that drove Reichenbach to propose his pragmatic justification of induction concerns a central feature of the frequency interpretation of the probability calculus. Familiarity with the details of the probability calculus will be assumed here, and with the fact that it is compatible with at least several interpretations. The axioms of the probability calculus are, of course, as follows:

(A.1) $P(a) \geq 0$ for all a in the domain of $P(\cdot)$.

(A.2) $P(t) = 1$ if t is a tautology.

(A.3) $P(a \vee b) = P(a) + P(b)$ if a and b and $a \vee b$ are all in the domain of $P(\cdot)$, and a and b are mutually exclusive.

Recall that on Reichenbach's frequency interpretation of probabilities such quantities are to be construed as tautological consequences of the probability calculus.¹⁰ More importantly, probabilities are to be regarded as measures of the limit of the relative frequency with which one contingent property is associated with another in an infinite sequence. More formally, the relative frequency of a pair of properties in a sequence is to be defined as follows:

$$F^n(A, B) = N^n(A, B) / N^n(A)$$

Here $F^n(A, B)$ is the frequency of associated As and Bs in a sequence of length n . Given this conception of relative frequency we can then define the concept of probability as follows:

$$P(A, B) = \lim_{n \rightarrow \infty} F^n(A, B).^{11}$$

Having introduced this notion of probability Reichenbach then proposes the rule of induction that states,

If an initial section of n elements of a sequence x_i is given, resulting in the frequency f^n , and if, furthermore, nothing is known about the probability of the second level for the occurrence of a certain limit p , we posit that the frequency f^i ($i > n$) will approach a limit p within $f^n \pm \delta$ when the sequence is continued (Reichenbach, 1949c, p. 47).

However, these definitions give rise to some very difficult but well-known problems concerning the existence of infinite sequences and the existence

¹⁰ See Reichenbach (1949b) and Weatherford (1982), chapter 4.

¹¹ See Reichenbach (1949c) for details concerning how this derivation is carried out.

of such convergent limits.¹² Reichenbach assumes that we are only ever aware of sequences that "...are not *intensionally* given, but are presented to us only by enumeration of their elements, i.e. are *extensionally* given (Reichenbach, 1949a, p. 309)" and that any such sequence of observed associations will be finite. Upon considering further extensional enumeration of the elements of a given observed sequence, we find that such extended sequences are, in point of fact, compatible with any value of the limit frequency. If this is so, we might ask why we are entitled in any way to assume that the frequency of such an association in even very long sequences of observed associations in a population will justify our assertion that that frequency will not diverge in further extensive enumerations of that sequence. Unless one is prepared to reject extensional empiricism this conclusion seems inescapable. Adopting intensionalism and accepting the existence of the relevant sort of a priori knowledge, like Russell ultimately did (1912 and 1948), permits induction to be grounded in a robust essentialism and a form of rationalism that allows for synthetic *a priori* knowledge of universals. But, although Reichenbach himself endorsed such a view in his early thinking about this problem in his dissertation, in his later work he staunchly opposed this sort of view and the idea of a priori knowledge, claiming that, "The idea that there is such a thing as a 'rational belief' is the root of all evil in the theory of knowledge and is nothing but a remnant from rationalistic philosophies (Reichenbach, 1949d)."

In any case, Reichenbach saw that his understanding of this problem in terms of frequentism was just the classical problem of induction in a somewhat new guise, and he showed two things. First, he showed that, by definition, *if such a limit exists*, then the procedure of scientific induction will be successful, and, second, that scientific induction is at least as good as any other method in discovering what is really the case concerning the frequency of an association in a sequence. Reichenbach explains,

Let us assume for the moment that there is a limit towards which the sequence converges, then there must be an n from which on our posit [the rule of induction] leads to the correct result; this follows from the definition of the limit, which requires that there be an n from which on the frequency remains within a given interval δ . If we were to adopt, on the contrary, the principle of always positing a limit outside $f^n \pm \delta$ when a frequency f^n has been observed, such a procedure would certainly lead us to a false result from a certain n on. This does not mean that there could not be other principles which like the first [the rule of induction] would lead to the correct limit. But we can make the following statement about these principles: even if they determine the posit outside $f^n \pm \delta$ for a smaller n , they must,

¹² Sequences with convergent limiting frequencies are just those sequences that settle into stable frequencies in the limit.

from a certain n on, determine the posit within $f^n \pm \delta$. All other principles of positing must converge asymptotically with the first [the rule of induction] (Reichenbach, 1949b, p. 316).

What he showed was that if a limit exists for a sequence, then by repeated application the rule of induction will lead to the value of that limit to any desired degree of approximation in a finite number of applications, and that all other methods will asymptotically converge with the results of the rule of induction. So, in spite of the fact that we cannot know that the limiting frequencies of sequences exist, we might as well simply accept the rule of induction because it is the best method of all methods. All methods are, in a sense, parasitic on the rule of induction. Again, this pragmatic answer to the problem of induction arose directly out of Reichenbach's recognition that, in point of fact, *we cannot know that such limits exist in our world*. We cannot know whether such convergent limits exist based on the empirical observation of associations in finite, extensionally given, sequences. So, we are stuck in the situation that either no method at all works, or induction is the best of all methods. In terms of the sophisticated Reichenbach argument this can be expressed as follows. If worlds are extensional and there is no a priori knowledge of the regularity of the world, then we cannot know the real value of $m(w_a)$. Nevertheless, it will be true that if $1 \geq m(w_a) > \varepsilon$, then scientific induction will be successful. If this is not the case, then no method will be successful.

3. Posits and Pragmatic Justifications

Even so, Reichenbach tells us that we can treat the existence of such limits of relative frequencies as *posits*, where posits are not to be treated as beliefs in the normal sense, but rather as a kind of wager concerning what would be most advantageous to us. Reichenbach explains that,

It is evidently the concept of posit which we have to employ for an explanation of this method. If in the finite section given we have observed a certain frequency f^n , we posit that sequence, on further continuation, will converge towards the limit f^n (more precisely: within the interval $f^n \pm \delta$). We posit this; we do not say that it is true, we only posit it in the same sense as the gambler lays a wager on the horse which he believes to be fastest. We perform an action which appears to us the most favorable one, without knowing anything about the success of this individual action (Reichenbach, 1949b, p. 315).

Furthermore, as all other rules are parasitic on the rule of induction, it is only natural to lay our wager on that rule. We are wagering that $1 \geq m(w_a) > \varepsilon$. So, the sort of justification his argument provides is clearly a matter of pragmatics. It is, as Feigl claimed, a justification actionis.¹³ But,

¹³ See Feigl (1963).

in any case, the kind of wager involved in positing the existence of convergent limits in infinite sequences is not the typical kind of wager that a gambler makes. Normally, a gambler at least knows the odds with which he is confronted and so can make an informed decision about what outcome to bet on (i.e. which is the best bet), but in the case of the limits of infinite sequences we are making the posit that the limit converges to f^n blindly; i.e. we are making this posit when we do not know the odds, and so we do not know if it is the best posit.

Reichenbach claims that in such cases we are making what he calls an *approximative posit* concerning the existence of such limits. We are blindly wagering that $1 \geq m(w_a) > \varepsilon$. As we have seen, Reichenbach shows that *if we are right about the existence of such a limit* (if this blind wager is correct), then induction will be successful, and if any other method is successful, then scientific induction will be successful. *If we are wrong about the existence of such a limit* (if this blind wager is not correct), then if any other method is successful, then scientific induction will be successful in this more restricted sense. Therefore, scientific induction is at least optimal in this specific sense. However, as BonJour notes in the passage quoted in section 1, this by no means shows that induction is justified in the traditional sense, and Reichenbach's view is apparently compatible with radical skepticism concerning the probity of induction. It may simply be false that $1 \geq m(w_a) > \varepsilon$ and, given extensional empiricism, we cannot know whether this claim is true or false. So, as far as we know, the method of induction might well be the best of a bad lot. Nonetheless, Reichenbach argues that there is a sense in which his argument *vindicates* induction if one is committed to extensional empiricism. It does show that if any method works, then induction works. We do not *know* that the rule is unreliable, but we know that it is the best method if any method is reliable. So, why not commit ourselves to the use of scientific induction? Of course, this may not be a satisfactory justification for someone who has sympathies with BonJour's inductive skeptic, but it is clearly to our practical advantage if scientific induction turns out to be reliable. As Reichenbach sees it, the alternative is to give up on science altogether. More importantly for the purposes of this discussion, what this result really establishes is that *given extensional empiricism* induction can *only* be pragmatically justified in the sense of Reichenbach's vindication. But how? Reichenbach is not terribly clear on this point, but this lacuna can be fixed easily by appeal to the decision-theoretic maximin rule.

4. The Maximin Justification of Reichenbach's Rule

After presenting his account of the optimality of the inductive rule, Reichenbach discusses the sense in which using this rule can be justified in

light of the problems associated with our ignorance of the existence of the convergent limit sequences that are required for the success of induction. As we have already seen, he says of the inductive rule that,

We posit this; we do not say that it is true, we only posit it in the same sense as the gambler lays a wager on the horse which he believes to be fastest. We perform an action which appears to us the most favorable one, without knowing anything about the success of this individual action (Reichenbach, 1949b, p. 315).

Elsewhere he also says:

Inductive positing in the sense of a trial-and-error method is justified so long as it is not known that the attempt is hopeless, that there is no limit of the frequency. Should we have no success, the positing was useless; but why not take our chance?

The phrase "take our chance" is not meant to state that there is a certain probability of success; it means only that there is a possibility of success in the sense that there is no proof that success is excluded (Reichenbach, 1949c, p. 49).

Finally, he explains to Russell in his 1949 letter that,

I have shown that blind posits are justified as a means to an end, and that no kind of belief in their truth is required. This I regard as an essential merit of my theory: I have shown that there are other reasons to make assertions than reasons based on belief (Reichenbach, 1949d, p. 407).

So, these remarks show that Reichenbach is clearly asserting that there are alternative justifications that have to do with non-epistemic pragmatic considerations and we know only that the probability that there are limit frequencies is non-zero. But, he also tells us in his 1949 letter that,

...the inductive conclusion can be called probable only when many other inductions have already been made, which tells us something about the second level probabilities. I speak here of the stage of *advanced knowledge*. In *primitive knowledge*, i.e. before any inductions were made, the inductive conclusion is not probable (Reichenbach, 1949d, p. 409).

So, in point of fact, Reichenbach believes that it is only when we have begun to make inductive inferences that we can even claim that the conclusions of inductive inferences are probable, and this is based on the success of making such inductions. Thus, the second-order probability of the inductive rule itself is very low in the state of primitive knowledge, for we have not therein yet made first-order inductions sufficient to gauge the second-order probability of induction. Prior to this state of advanced knowledge, inductive conclusions are not probable and can only be justified pragmatically in the sense of blind positing. Notice that this will be true for *any* alternative inductive rule as well. So, the rationality of induction writ large and independent of the selection of one of the possible reliable inductive rules depends on their being a vindication of the inductive rule in this primitive state. Even if there are other rules that are superior to Reichenbach's inductive rule in terms of how often they get correct results,

this can only be determined in the state of advanced knowledge. All of this then suggests that Reichenbach's rather vague appeal to pragmatic justification can be made much more coherent by appealing to some particular resources in modern decision theory and focusing on justifying induction in the primitive state of knowledge.

In decision situations where probabilities cannot be meaningfully assigned to outcomes (other than that they are non-zero) advice about what to do is wholly a function of utilities, because no expected utilities can be calculated for such cases. There are then two ways one might provide such a pragmatic justification for adopting the inductive rule. First, in such cases and where the potential losses are great Wald's maximin rule usefully applies:

(MR) maximize the minimum utility outcome.¹⁴

Reichenbach's approximative posit fits this bill perfectly. Much is at stake here. If scientific induction does not work, then our inductive inferences cannot possibly succeed and we would have to give up on science. Moreover, as Reichenbach admits, we have no idea of the probabilities on which this wager is based. So, where R is ' $1 \geq m(w_a) > \epsilon$ ', I is 'employ scientific induction', S is 'assured success in inductive reasoning in the long run', P is 'possible success in the long run', Reichenbach's wager can be fruitfully understood as follows:¹⁵

	R	$\neg R$
I	S	$\neg S$
$\neg I$	P	$\neg S$

¹⁴ See Wald (1939, 1945) and Resnik (1987, pp. 26-27).

¹⁵ Since we are only considering the primitive state of knowledge here, we need not be concerned with the claim that there are alternative rules of induction that are superior to Reichenbach's preferred rule in our table. We can simply see that whatever rule in particular we adopt as the second-order guide to inductive inference, Reichenbach's considerations show us that the only kind of rational justification for any inductive rule (no matter how efficient) in the primitive state is a blind posit of this pragmatic sort. We could, in effect, run the very same argument substituting for I any such second-order rule. So, the upshot of Reichenbach's considerations really should be that the only justification for induction in the primitive state of knowledge is pragmatic.

Here we also have the following relative ordering of utilities: $V(S) > V(P) > V(\neg S)$, and this is all the application of the maximin rule requires. We do not need to know ordinal utilities to apply MR. MR tells us to look at the worst outcomes of the acts given the relevant states of the world. For $\neg I$ we get $\neg S$, for I we also get $\neg S$. These values are equal. By the lexical MR we then look at the next lowest outcome(s) of $\neg I$ and I respectively. In the case of I we get S , and for $\neg I$ we get P . So, the MR tells us that it is maximally rational to do I even in the primitive state of knowledge where inductive conclusions are not probable and where the second-order probability that the inductive rule is correct is also very low.

The second way one might flesh out Reichenbach's solution is in terms of a simple dominance argument based on the dominance principle. Again, where we have ignorance of the relevant probabilities we should look at the utilities of the outcomes and apply a principle of pragmatic rationality to decide what to do. The dominance principle is just such a rule. The familiar notion here is that an act A dominates an act B if for every outcome the utility of A is equal to or greater than the utility of B and for at least one outcome the utility of A is greater than that of B . The dominance principle is then this:

(DOM) opt for dominant acts.

Referring back to our decision table we can see that I dominates $\neg I$ even in the primitive state of knowledge where inductive conclusions and the second-order probability that induction is correct is very low. Nevertheless, DOM tells us that it is pragmatically rational to opt for I . Thus, Reichenbach's solution can be pragmatically vindicated, but in a principled rather than intuitive manner. It *is* pragmatically rational to use scientific induction and there is nothing at all obscure about why this is so. It is pragmatically justified in light of the maximin rule and/or in light of the dominance principle, and this is so despite our ignorance of the relevant probabilities concerning the existence/non-existence of convergent limits, the primitive improbability of inductive conclusions, and the improbability of the second-order claim that the inductive rule is correct.

References

- BonJour, L. (1986). A Reconsideration of the Problem of Induction. *Philosophical Topics*, 14, 93-124.
- BonJour, L. (1992). Problems of Induction. In E. Sosa and J. Dancy (eds.), *A Companion to Epistemology* (pp. 391-395). Cambridge: Blackwell.
- BonJour, L. (1998). *In Defense of Pure Reason*. Cambridge: Cambridge University Press.
- Eberhardt, F. (2011). Reliability via Synthetic A Priori: Reichenbach's Doctoral Thesis on Probability. *Synthese*, 181, 125-136.
- Feigl, H. (1963). De Principiis Non Diputandum...? In M. Black (ed.), *Philosophical Analysis: A Collection of Essays*, (pp. 113-131). Englewood Cliffs: Prentice Hall.
- Galavotti, M. (2011). On Hans Reichenbach's Inductivism. *Synthese*, 181, 95-111.
- Hájek, A. (1997). 'Mises Redux' — Redux. Fifteen Arguments against Finite Frequentism. *Erkenntnis*, 45, 209—227.
- Hájek, A. (2009). Fifteen Arguments against Hypothetical Frequentism. *Erkenntnis*, 70, 211—235.
- Kelly, K. (1991). Reichenbach, Induction and Discovery. *Erkenntnis*, 35, 123-149.
- Kelly, K. (1996). *The Logic of Reliable Inquiry*. Oxford: Oxford University Press.
- Reichenbach, H. (1949a). On the Justification of Induction. In H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis* (pp. 324-329). New York: Appleton-Century-Crofts.
- Reichenbach, H. (1932/1949b). The Logical Foundations of the Concept of Probability. In H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis* (pp. 305-323). New York: Appleton-Century-Crofts. Originally published in 1932 as Die Logischen Grundlagen des Wahrscheinlichkeitsrechnung. *Erkenntnis*, 3, 410-425.
- Reichenbach, H. (1949c). *Theory of Probability*. Berkeley: University of California Press.
- Reichenbach, H. (1949d). A Letter to Bertrand Russell. In Reichenbach 1978, 405-411.
- Reichenbach, H. (1978). *Selected Writings, 1909-1953 (vol II)*. M. Reichenbach and W. C. Salmon (eds.). Boston: Dordrecht.
- Resnik, M. (1987). *Choices*. Minneapolis: University of Minnesota Press.
- Rosenkrantz, R. (1981). *Foundations and Applications of Inductive Probability*. Atascadero: Ridgeview Publishing.
- Russell, B. (1911). On the Relations of Universals and Particulars. *Proceedings of the Aristotelian Society*, 12, 1-24.

- Russell, B. (1912). *The Problems of Philosophy*. New York: Oxford University Press.
- Russell, B. (1948). *Human Knowledge: Its Scope and Limits*. New York: Clarion.
- Salmon, W. (1963). On Vindicating Induction. In Sidney Luckenbach (ed.) *Probabilities, Problems and Paradoxes* (pp. 200-212). Encino: Dickenson.
- Salmon, W. (1966). *The Foundations of Scientific Inference*. Oxford: Oxford University Press.
- Skyrms, B. (1966). *Choice and Chance: An Introduction to Inductive Logic*. Belmont: Dickenson Publishing Co., Inc.
- Teng, C. M. and H. Kyburg (2001). *Uncertain Inference*. Cambridge: Cambridge University Press.
- Wald, A. (1939). Contributions to the Theory of Statistical Estimation and Testing Hypotheses. *The Annals of Mathematics*, 10, 299-326.
- Wald, A. (1945). Statistical Decision Functions Which Minimize the Maximum Risk. *The Annals of Mathematics*, 46, 265-280.
- Weatherford, R. (1982). *Philosophical Foundations of Probability Theory*. London: Routledge.

From ‘Syntax’ to ‘Semantik’ — Carnap’s Inferentialism and Its Prospects*

Adam Tamas Tuboly

(a) *Institute of Philosophy, Hungarian Academy of Sciences*

(b) *University of Pécs*

Abstract: The aim of this paper is to provide context for and historical exegesis of Carnap’s alleged move from syntax to semantics. The Orthodox Received View states that there was a radical break, while the Unorthodox Received View holds that Carnap’s syntactical period already had many significant semantical elements. I will argue that both of them are partly right, both of them contain a kernel of truth: it is true that Carnap’s semantical period started after his *Logical Syntax of Language* — in one sense of semantics. But it is also true that Carnap had already included semantical ideas in LSL: though not (just) in the sense that URV maintains. This latter sense of semantics is related to what is usually called *inferentialism*, and by getting a clearer picture of Carnap’s original aims, context, and concept-usage, we might be in a better position to approach his alleged inferentialism.

1. Introduction: The Scene and the Heroes

Carnap held many views, though still not as many as Bertrand Russell — nonetheless, orthodox views about these changes emerged and stuck around during the last quarter of the twentieth century. One of them maintained that there was a radical break between Carnap’s syntactical and semantic period, the latter being initiated alone by Alfred Tarski. As A. J. Ayer (1982, p. 52) said, “Carnap’s eyes were dramatically opened at a congress in Paris in

* This paper was supported by the MTA BTK Lendület *Morals and Science* Research Group; and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. I am indebted to Jan Woleński and the Carnap Archives at Los Angeles (Rudolf Carnap papers (Collection 1029). UCLA Library Special Collections, Charles E. Young Research Library) and at Pittsburgh (Rudolf Carnap Papers, 1905-1970, ASP.1974.01, Special Collections Department, University of Pittsburgh) for the permission to quote the archive materials. All rights reserved. All the translations from the archive files are mine. I cite the Pittsburgh Archive files as follows: RC XX-YY-ZZ where XX is the box number, YY the folder number, and ZZ the item number; the UCLA archive as Carnap 1957, [UCLA] followed by box, folder, and page numbers. All the translations from the German letters are mine; the original, often idiosyncratic and ungrammatical English sentences of Carnap and Neurath’s letters are preserved.

1935 when Tarski presented an abstract of his semantic theory of truth.” This view may be labeled as the *Orthodox Received View* (ORV) of Carnap’s philosophical development.

It was pointed out recently that this picture is highly oversimplified; furthermore, certain parts of it are entirely false. These new developments — reaching their peak in the 1990s — seem to be the new orthodoxy, thus I will refer to them as the *Unorthodox Received View* (URV). I will argue that both of the received views are partly right, that is, both of them contain a kernel of truth: it is true that Carnap’s (1937) semantical period started after his *Logical Syntax of Language* (LSL; originally in German, 1934) — in one sense of semantics. But it is also true that Carnap had already included semantical ideas in LSL: though not (just) in the sense that URV maintains. This latter sense of semantics is related to what is usually called *inferentialism*, and by getting a clearer picture of Carnap’s original aims, context, and concept-usage, we might be in a better position to approach his alleged inferentialism.

One more bit of context: recently Jaroslav Peregrin (2014 and forthcoming) has uncovered the logical side of Carnap’s LSL and its relation to contemporary inferentialism. As he (forthcoming) argues, “from the current vantage point [...] Carnap’s investigations from LSL are much more interesting than even the later Carnap himself would have appreciated.” The recent paper is not so much about the logical technicalities and subtleties of Carnap’s work, but about the historical settings: some more context and historical evidence will be provided for the recent understanding of Carnap.

2. The Orthodox Received View

ORV has stuck around until the late 1980s, though one might still encounter it in textbook-type presentations. The main idea behind ORV is that Carnap’s characteristic ideas are either flawed, or turned out to be wrong. There are four main points that seem to be the core of this view and they could be made explicit as follows:

(A) During the early 1930s, Carnap embraced the approach of ‘logical syntax’ that holds that every philosophical problem is a linguistic (syntactical) problem, and it deals exclusively with formal structures and syntactical categories, without taking into account any type of meaning and/or reference. “Philosophy is to be replaced by the logic of science,” said Carnap (1937, p. xiii), “that is to say, by the logical analysis of the concepts and sentences of the sciences, for the logic of science is nothing other than the logical syntax of the language of science” (emphasis in the original). In the year of the *Syntax* book, Carnap (1934a, p. 29) even

declared that he and his Viennese friends “*pursue Logical Analysis, but no philosophy*” (emphasis in the original).

(B) The conception presented in *Logical Syntax* is full of serious philosophical and technical mistakes and errors; therefore the approach based on logico-syntactical analyses of language gives us a highly defective tool: it was doomed to failure.

(C) There is a radical break in Carnap's development: he was convinced by Alfred Tarski to give up his syntactical approach in favor of semantics even before the ink was dry on the *Syntax*-book.

(D) Finally, Carnap accepted the semantic method, and quickly worked out various semantical systems, using such notions as ‘truth’, ‘truth-condition’, ‘reference’, ‘denotation.’

The Orthodox Received View was eagerly defended and propagated, from different perspectives, by A. J. Ayer (1982), Max Black (1945, p. 172), Ernest Nagel (1942, p. 469), Maria Kokoszyńska (1936 and 1939, p. 118) and Otto Neurath. Some of them thought that Carnap was guilty of accepting semantics despite its obvious *unempiricist* and *metaphysical* underpinnings and/or implications. For example, Black (1945, p. 172) said that “[i]t comes as something of a shock [...] to find these new studies re-admitting wholesale those ostensible referents of symbolism which had seemed to have been permanently banished from ‘scientific philosophy’ [in LSL].” The American logical empiricist Ernest Nagel (1935, p. 357) also admired LSL, and thus raised in his review of *Introduction to Semantics* two important problems: first, it seemed that Carnap's approach to designation leads to unnecessary and problematic metaphysical hypostatization, and secondly that “it appears doubtful [...] whether a semantics which deliberately abstracts from all reference to the users of a language has much to contribute to the resolution of general philosophical issues” (Nagel, 1942, p. 472).

Even before Nagel's review of the book, Carnap was aware of his dismissal of the new semantical project. In *Introduction to Semantics* (1942, p. xi), Carnap stated that “while many philosophers today urge the construction of a system of semantics, others, especially among my fellow empiricists, are rather skeptical.” In the Neurath-Carnap correspondence, this is explained as follows: “When I wrote about fellow empiricists making objections against Semantics (in the Preface of vol. I) I was, of course, thinking in the first place of you and Nagel.”¹ Carnap, as he claimed in the same letter, did not name or quote them because they did not publish anything that was explicitly against semantics.

Nagel planned, however, to publish a paper on Charles S. Peirce around 1939: he aimed to attack Carnap's way of pursuing semantics, but

¹ Carnap to Neurath, May 11, 1943 (RC 102-55-01).

it did not appear since *The Journal of Unified Science* (the follow-up of *Erkenntnis*) was cancelled during the war.² Nonetheless, since the paper was read at the 1939 International Congress for the Unity of Science (at Harvard), Carnap was aware of its content. Nagel (1939/1954, p. 96) emphasized Peirce's behavioristic treatment of language: "neither terms nor statements can be regarded as *designating*, independently of the habits involved in their use" (emphasis in the original). Besides reconstructing Peirce's philosophy, Nagel accepted and recommended it with regard to the newly emerged semantics of Carnap as well; there is a long passage about how Nagel incorporated his fears into the narrative of the history of logical empiricism, and since it is quite unknown, I will quote it at full length:

Some have suspected, perhaps unjustly, that the recently inaugurated discipline of semantics will open wide the door for the rehabilitation of Bolzano's *Saetze-an-Sich*, Meinong's *objectives*, Russell's *subsistens*, and allied conceptions of the referends of signs. Though such doctrines have had fruitful historical roles, I think it would be a retrograde step if modern logical empiricism were to revive them in a new form; for the greater strength and promise of the movement has been its interpretation of the abstract in terms of the concrete, and its resolute turning from speculations which have no ascertainable consequences in issues of observable fact. (Nagel, 1939/1954, p. 96. Emphases in the original.)

Carnap (1942, pp. xi-xii) indeed knew about these fears and was inclined to dissolve them. He even referred to Nagel's talk implicitly: friends of the movement, he said, "are afraid that a discussion of propositions [...] and truth [...] will open the back door to speculative metaphysics, which was put out at the front door."

In January 1950, Carnap wrote a letter to his friend Franz Roh, an art historian from Munich: in that letter he happily accepted Roh's 'psychological and sociological explanation of misunderstandings' and used examples from his own life to confirm the narrative about how his old friends tried to save him from the new changes and conversions. "[Take] for example my transition from the period of my 'Logischen Syntax der Sprache' to the later semantical period to which my latest books belong," wrote Carnap to Roh, "[...] my friends always came and said, 'But you have put everything such a good way; you cannot suddenly overthrow everything!'"³ Because so many of his friends were criticizing him for the propagation of semantics and they did not understand that it was a practical decision and not a theoretical question, he wrote up a paper to explain his stance: it was "Empiricism, Semantics, and Ontology" (ibid.).

² Nagel's paper was published later in his selected essays: *Sovereign Reason and Other Studies in the Philosophy of Science*; see Nagel (1939/1954).

³ Carnap to Franz Roh, January 14, 1950 (RC 102-34-06).

Thus the targets of Carnap's famous paper were Nagel and Neurath, and not just the usual suspect W.V.O. Quine.

Though the above passage might be seen as a digression, actually it documents well Carnap's basic attitude towards philosophy and his friends: make everything as clear as it can be and do not let misunderstanding influence your judgment and make you look for theoretical gaps and clashes when only practical matters surface. The question of the nature and status of semantics was for Carnap such a milestone that it deserved much more attention.

From the Nagel-Neurath pair, the latter presented a more difficult case: it was quite difficult to untangle Neurath's arguments against semantics, as they amalgamated the requirement of clarity for empirical languages, a vehement anti-metaphysical attitude, and certain considerations against non-pluralistic (totalitarian) tendencies in science and especially in politics. While LSL was Neurath's "logical Bible,"⁴ *Introduction to Semantics* presented just "all the Aristotelian metaphysics in full glint and glamour," and the same absolutist attitude that might be found (at least Neurath detected it) behind all those searches for the *real, one world*, for *the only Truth* that characterized the entire tradition from Plato, through Descartes and Kant, to German idealism and finally to fascism and Nazism.⁵

Others, mainly the Polish logicians, maintained that the earlier Carnap was a bad guy. Perhaps due to either being under the wrong influences, or working in a sort of isolation, Carnap promoted mistaken and false ideas in his syntactical phase, and it took "a Tarski" to enter the picture and single-handedly transform him into an advocate of semantics. This narrative, though it is diametrically opposed to Neurath's and Nagel's story, provided its own 'radical break' point in Carnap's philosophical and intellectual development.

Be as it may — I am not here to judge which Carnap was "the true and right Carnap" — both groups (and sometimes even Carnap!) claimed explicitly or implicitly that there was a break in his intellectual development, centered on the idea of semantics, initiated by Polish logicians, or by one particular logician, namely Tarski.

3. The Unorthodox Received View

⁴ Neurath to Carnap, November 25, 1934 (RC 029-10-07).

⁵ Neurath to Carnap, January 15, 1943 (RC 102-55-02). On semantics and politics see Neurath to Carnap, September 15, 1943 (RC 102-55-03), this letter contains the term "Plato-Hitler") and September 22, 1945 (RC 115-07-66, unsent). Cf. Reisch, 2005.

Beginning with the works of Alberto Coffa (1976; 1991), Thomas Ricketts (1996), Stewe Awodey (2007), and Richard Creath (1990a; 1996; 1999), a new understanding of Carnap emerged. Creath (1996, p. 251) argued that “for the most part, [...] Carnap’s syntax is really semantics or at least that the difference between what Carnap provides and a full-blown semantics would be infinitesimal were there infinitesimals.” Their points were the following:

(1*) Even though Carnap argued for a *syntactical* conception of language, logic, and philosophy, his project involved many elements and techniques which later came to be called *semantics*. Creath summarized (1990a, p. 410) this as follows: “Not only is the logical consequence relation itself semantical, as we use the term, but so are truth tables, interpretation, and analyticity, all of which Carnap discusses. The treatment, especially of the last of these, is surprisingly close to a full semantical account.” Even Tarski noted in his “On the Concept of Logical Consequence” (1936/1956, p. 413) and in the “Historical Notes” for “The Concept of Truth in Formalized Languages” (1935/1956, pp. 277–278) that Carnap achieved very similar results via very similar methods regarding truth and semantics.

(2*) Carnap’s main metaphilosophical commitments (conventionalism, holism, or in Creath’s terms, functionalism) remained intact in the alleged change of views, indicating that his radical break was not *that* radical. Carnap suggested a pragmatic approach to logic, syntax, and semantics in his later works, as well as in his famous principle of tolerance in LSL (1937, §17). He claimed that semantics is to be regarded as a *tool* that is used to structure and systematize knowledge, and its use should be decided by its *expedience* (Carnap, 1950/1956, p. 221). It would be important to mention that there were scholars already in 1942, who claimed that — from a general point of view — “the continuity with Carnap’s previous thinking is evident from the beginning” since he ignores “psychologistic pseudo-problems” (J.M.R., 1942, p. 282).⁶

(3*) Therefore, given the metaphilosophical continuation and the presence of semantical-technical matters in LSL, the Unorthodox Received View holds that the change between Carnap’s syntactical and semantical period is a relatively small one, and not a radical break.

⁶ Actually, Carus (1999) claimed that there isn’t any radical break between the syntactical and the semantical Carnap — or at least surely not in the sense of ORV — since all those conceptions, like theories of truth, coherence, correspondence, that are used by ORV to mark the gaps in Carnap’s thinking were rejected by Carnap entirely.

Ricketts (1996, p. 231) even claimed, “the step is so small that we should ask why Carnap needed the impetus of Tarski’s work to take it.”⁷

I tagged this approach as “Unorthodox Received View” because even though the aforementioned scholars made many new, important, and revealing discoveries and remarks, calling our attention to a much more interesting picture and context of Carnap’s work, their views have apparently become the new orthodoxy. Nevertheless, there is *still much more* to Carnap’s syntactical and semantical period, and using a bit more context and materials, we might be able to show (at least as a first approximation) that we still have to refine our new Carnap-picture.

4. A Timeline: From ‘Metalogik’, through ‘Semantik’, to ‘Syntax’

In this section, in order to solve some conceptual issues and to see how Carnap’s systems and ideas evolved, I will discuss three notions that he used frequently in the early 1930s, namely ‘Metalogik,’ ‘Semantik’ and ‘Syntax.’

When dealing with the Wittgensteinian problem of how to account for and talk about language, Carnap (1963, pp. 53-54) had a famous “sleepless night” in January 1931, when he formulated his new project: “Versuch eine Metalogik,” i.e. “Attempt at a Metalogic.” In his first Vienna Circle lectures on metalogic, delivered in June, he defined the notion as “the theory of the forms which appear in a language, thus the representation of the syntax of language. In it one must not — to follow the formulation of the Warsaw group — make reference [Bezug] to the meanings [Bedeutung] of the signs” (Stadler, 2001/2015, p. 107).⁸

Thus “Metalogik” captured Carnap’s intention: he wanted to talk about language without considering the various non-verbal elements; however, since language is determined by logic, the first step was to talk about *logic*, to work out a *metalogical* framework. During 1931, Carnap was thinking about his new project under the title and category of “Metalogik.” This might be seen also in his famous “The Elimination of Metaphysics Through Logical Analysis of Language” paper: he talks there about the “metalogical” formulations of language (1932/1959, p. 62). By the time of

⁷ It should be mentioned that Tarski did not ‘open Carnap’s eye dramatically’ in 1935, since it is known that Carnap invited Tarski to Paris, and that he was involved in the German translation of Tarski’s 1933 truth paper. On this, see Gruber (2016) and Woleński (2017).

⁸ Carnap’s three lectures on metalogic, introducing his new ideas to the Circle, have been translated and published in Stadler (2001/2015, pp. 107-111 (first lecture), pp. 111-116 (second lecture), pp. 116-123 (third lecture and discussion)).

the publication of this paper, Carnap wrote from Prague to Schlick that “of my ‘Metalogic’, the main part is done.”⁹

Even though the title of the manuscript was still “*Metalogik — Die logische Syntax der Sprache*” in 1932, important changes were lurking around. He reported to Schlick in June that he “accepted the suggestion of Gödel and Behmann to replace ‘Metalogic’ with the word ‘Semantics.’”¹⁰ It is known from Carnap’s correspondence that he asked Gödel (January 1) to read the manuscript and make a “few critical remarks.”¹¹ Gödel was interested and Carnap sent him via Hempel the “*Metalogik*” manuscript (February 23, 1932; Gödel, 2003, p. 343). Their letters dealt with Carnap’s definition of “analyticity,” but nothing was said about Syntax/Semantics, only Carnap’s diaries testify that Gödel suggested the term “*Semantik*”; actually it happened on the same day when Neurath proposed “*Universal Syntax, logical foundations of Unified Science*” (“*Universelle Syntax, logische Grundlagen der Einheitswissenschaft*”) as a new title to emphasize connections to the movement.¹²

To see why Gödel might have suggested the term “*Semantik*,” another player shall enter the scene now. The German mathematician and logician Heinrich Behmann also got the manuscript before January 1932: that time Carnap asked him terminological questions, for example, about the title: “Is ‘Metalogic’ good? or is ‘logical’ Syntax better?”¹³ Behmann read the manuscript, made many technical remarks that will not concern us here, and responded to Carnap’s title-questions. Interestingly, Behmann indicated that, the problem was *more* than just a *terminological* issue. He claimed on March 23 that “*metalogic*” is too narrow a term, since what captures Carnap’s intentions is not an investigation about logic, as the term would suggest, but about language, so analogously a term like “*Meta-Sprache*,” “*Meta-language*” would be a better fit.¹⁴ However, three days later (March 26), Behmann added some notes to the same letter; he came up with a new and better suggestion: what actually captures Carnap’s new project is “‘Semantics, as ‘the study of representation [*Darstellung*] through signs’” (*ibid.*).

The term “*Semantik*” stuck with Carnap: in the two lectures that he prepared for Hans Reichenbach’s Berlin-colloquium in July 1932, he talked about the “*Formal Questions of Semantics*” and “*Semantics as the*

⁹ Carnap to Schlick, December 7, 1931 (RC 029-029-15).

¹⁰ Carnap to Schlick, June 30, 1932 (RC 029-029-11).

¹¹ Carnap to Gödel, January 1, 1932 (RC 115-08-06).

¹² Carnap’s diary entry and margin note, March 26, 1932 (RC 025-75-10).

¹³ Carnap to Behmann, January 20, 1932 (RC 115-10-16).

¹⁴ Behmann to Carnap, March 21, 1932 (RC 115-10-18).

Foundation of Scientific Philosophy.”¹⁵ In the second lecture, we read the following:

What is semantics? Hilbert's metamathematics or proof-theory [Beweistheorie]: signs as objects = semantics of mathematical systems of formulas.

Semantics = theory of forms of the complete language [Formtheorie der Gesamtsprache], combinatorics of the sign-complexes, without reference [Bezugnahme] to the referents [Bedeutung] of signs! (Like chess). Calculus. 2 parts: 1) Syntax, 2) theory of deduction [Schlusslehre] [...]

Thesis: Scientific Philosophy is Semantics. [...]

[...]

2) Philosophy handles the language-forms, without reference to the referents!

Example: Consequence-relation [Folgebeziehung], the most important concept of Semantics. (RC110-07-23, p. 1. Underlining in the original.)

Even though Carnap crossed “Semantik” in the titles and wrote “Syntax” everywhere, there was apparently a point when he was more than positive about semantics: it was just the foundation of scientific philosophy! This period, however, might be characterized as “oscillation”: in the reply to his fellow Vienna Circle member Edgar Zilsel and to the Gestalt psychologist Karl Duncker, written mainly during the summer of 1932, Carnap (1932, p. 177) mentioned semantics, logical syntax, and metalogic at the same time, since *all* of them study the *structure* of sentences of a given language.

At this time ‘metalogic’ was the least favored term and Carnap was caught between ‘syntax’ and ‘semantics.’ In April 1932, he replied to Behmann’s letter as follows: “I like the term ‘semantics’; Gödel also suggested it. Neurath, on the other hand, thinks it is unappealing and pedantic; he suggests ‘syntax.’ In order to avoid confusion with syntax in philology, one would probably often have to call it ‘logical syntax.’”¹⁶ He wrote the same to Schlick: “I replaced the word ‘metalogic’ with ‘semantics’, accepting Gödel’s and Behmann’s suggestion.”¹⁷ Furthermore, in the table of contents, written for an earlier version of the *Syntax* book, Carnap had the following handwritten note at the title (which was at that time “Metalogik: Die logische Syntax der Sprache”): “Later ‘metalogic’ was replaced by ‘semantics’, but since Neurath rejected the word ‘semantics’, I kept ‘syntax’” (RC 110-04-07, p.1).

Nevertheless, until the end of 1932, Carnap referred to his manuscript (almost) always as ‘Semantik’ in his diaries. Though in February 1933 he

¹⁵ “Formale Fragen der ~~Semantik~~ Syntax (Für Vortrag in Reichenbachs Seminar, Berlin, 1.7.32.” (RC 110-07-20); and “Die ~~Semantik~~ Syntax als Grundlage der wissenschaftlichen Philosophie. Referat in Reichenbachs Kolloquium, Berlin, 4.7.32.” (RC 110-07-23).

¹⁶ Carnap to Behmann, April 17, 1932 (RC 115-10-19).

¹⁷ Carnap to Schlick, June 30, 1932 (RC 029-29-11).

wrote to Quine that during his “Logik II” course in Prague he “will discuss [the] new research on logical syntax (metalogic, semantics),”¹⁸ the primary category of identification became ‘syntax’ after that spring. Actually, when Carnap delivered a talk at the meeting of the Vienna Circle (July 10, 1933), he talked only about syntax and argued that “(philosophy =) logic of science = Syntax” (RC 110-07-22). After these lectures — where Carnap presented the famous “principle of tolerance” for the very first time — the main ideas of LSL were settled just as we know them today.

Before I move on, one more thing should be mentioned. In his intellectual autobiography, Carnap (1963, p. 54) remembered the discussed changes as follows: “At that time I defined the term ‘metalogic’ as the theory of the forms of the expressions of a language. Later I used the term ‘syntax’ instead of ‘metalogic’ [...]” ‘Semantics’ is obviously and notably missing from these remarks — thus vindicating the narrative of ORW. The original and unpublished autobiography, however, is a bit different: “At that time I defined the term ‘metalogic’ as ‘the theory of the forms of the expressions of a language, that is, the description of the syntax of language’ [...]. Later I used instead of ‘metalogic’ the term ‘~~semantics~~’; but I soon abandoned it in favor of ‘syntax’ [...]”¹⁹ Thus vindicating URV. It would be interesting to hypostasize about why Carnap deleted that clause, but that may lead too far into the history of American philosophy of science and language.

What conclusions shall we draw? (1) Though Carnap had various *names* for his project, *all of them captured something*. ‘Metalogik’ captured the idea that language is to be accounted and approached from a logical point of view. ‘Semantik’ was apt because it was connected to the idea (as we will see in the following sections) that Carnap wanted to talk about meaning and content, and these were semantical notions. Finally, ‘Syntax’ captured the idea that philosophy deals only with structures, formal rules and not extra linguistic entities. The different names thus documented the changes and development of Carnap’s perspectives and emphases.

5. “Semantik” in the 1920s and 1930s

Around the *fin-de-siècle* and shortly after, many words and concepts were in use to mark the studies of language, meaning, and reference: semasiology, rhematic, sematology, glossology, comparative ideology, sensifics, significs, rhematology, semiotic, semiology, orthology, and science of idiom (Read, 1948; Burgess, 2008 and ms.; Woleński, 1999).

¹⁸ Carnap to Quine, February 6, 1933; quoted from Creath (1990b, p. 109).

¹⁹ Carnap, Rudolf 1957 [UCLA]: “Autobiography. Part Two. Philosophical Problems.” Box 2, CM3: M-A4.

The application of some of these became regular only in the second half of the 19th century and they covered the investigations of the historical changes of word-meaning as an extension of the etymological method, or the method of dealing with sentences and how they function as signs in communication. So the study of language had many forms and various conceptions in actual usage, but none of them formed a detailed research program or a discipline among philosophers in continental Europe.

"Semantik" and closely related terms came to be used widely only relatively late in the 1920s, mainly by Polish logicians and philosophers. As Jan Woleński (1999 and 2009) pointed out, there were two main trends in the Lvov-Warsaw School. One of them originated from and was influenced by the works of Kazimierz Twardowski, a pupil of Franz Brentano. Among his most important students and colleagues, we find Jan Łukasiewicz, Stanisław Leśniewski, and Tarski himself. All of them understood semantics as a study of the word-world relation: "[P]olish philosophers, influenced by Brentano via Twardowski," said Woleński (2009, p. 52), "transformed the intentional conception of mental acts into a referential treatment of language." Interestingly, Neurath wrote the following to Carnap in January 1943:

Scholasticism created Brentanotism, Brentano beget Twardowski, Twardowski beget Kotarbinski, Lukasiewicz (you know his direct relations to the Neo-Scholasticism in Poland), both together beget now TARSKI etc and now they are God fathers of OUR Carnap too, in this way THOMAS AQUINAS enters from another door Chicago, where he entered already via [Mortimer] ADLER. (Neurath to Carnap, January 15, 1943. RC 105-55-02.)

For Neurath, this line of heritage represented such a grandiose metaphysical inheritance that it threatened logical empiricism with pseudo-scientific system building. He expressed his fears in the context of semantics and the new correspondence-like theory of truth claiming that a new "METAPHYSICA MODO LOGISTICA DEMONSTRATA," that is, a new metaphysics, based on strict logic, is hanging over the movement's head, threatening even its elementary integrity. But that is another story.²⁰

The referential treatment of language is clearly seen in Tarski's ideas: for him, the semantic method involved such concepts as denotation, satisfaction, definition, and truth. Most of these items were simply neglected at the time by those English and French-speaking scholars who tagged their projects with the above-mentioned categories of semantics. John Burgess (ms.) even claimed that "Tarski could have called what he actually calls 'semantic' notions 'referential' notions instead." It is quite interesting that he excluded almost all talk of meanings from his semantics.

²⁰ See Neurath to Carnap, July 1, 1936 (RC 102-52-23) and April 30, 1935 (RC 029-09-57).

One cannot find the term *synonym*, *homonym*, or *ambiguity* in Tarski, notes Burgess, “which surely would all have been near the top of a list of typical ‘semantic’ notions for [Michel] Bréal,” who was the father of semantics with his lectures and book at the end of the 19th century.

In 1945, the famous lexicographer and scholar of English slang, Eric Partridge complained on the pages of the *Journal of the Royal Society of Arts* that the concept of semantics, developed by Bréal and others was

[...] not good enough for Professor Carnap and his Polish, Austrian and American fellow logicians. They must adopt it, as though no other designation could have been found to suit their non-philological, their psychological, especially their logistical (rather than logical) purposes, and adapt it to their non-linguistic ends. For them, the word [semantics] means something quite different from the denotation intended and explained by Bréal, Jespersen, and all other true philologists. (Partridge, 1945, p. 455.)

So, obviously, not everybody was happy with the transformation of the old science of semantics — studying communication, information-production, signifying etc. — into the formal, mathematical and logical study of an abstract word-world relation. Carnap even noted in his unpublished autobiography that after the first 1935 Paris congress a reporter wrote about Tarski’s lecture as follows: “A learned philosopher came all the way from Warsaw to Paris in order to give to us the great revelation that snow is white if and only if snow is white.”²¹

Let’s stop for a minor and short digression. As can be seen from the remarks above, the opinion on logical semantics among professionals and laymen varied from acceptance to rejection and incomprehension. Another good example is the American economist, social theorist and popular writer Stuart Chase who published his *Tyranny of Words* in 1938. Propagating Alfred Korzybski’s theory of semantics, Chase (1938, p. 160) mentioned at one point that the University of Chicago Press will publish a new encyclopedia under the leadership of Otto Neurath, who “is a kind of pioneer in semantics. He believes in going to things wherever possible, rather than to words.” Neurath wrote to Carnap that after he moved to England, he rebuilt his library, which now included many new books, among them items from Stuart Chase.²² It is unknown whether *Tyranny of Words* was among them, or if it was, whether Neurath ever read it carefully (though the title might have caught his attention): what is sure, however, is that Neurath would never accept that he is a pioneer in semantics, or that he believed in going to things and not to words — or at least not without major restrictions and qualifications.

²¹ Carnap 1957 [UCLA], Box 2, CM-3: MA-4, M7-M8.

²² Neurath to Carnap, July 17, 1942, (RC 102-56-04).

Thus it does not seem unreasonable that Carnap did not want to go public with the debate concerning semantics. In his opinion, it would not have helped their common cause. As he wrote to Neurath, "I think for the sake of the movement it would be much better if we were more tolerant towards each other. If your intolerance would become the general custom, then I am afraid you would be among the first to be declared a heretic and excommunicated."²³ The major problem was that they couldn't even agree on what exactly they disagreed about. In Neurath's narrative they had a theoretical debate with some political and social consequences; in Carnap's eyes, however, they had some practical issues and he thought that "it would be best for the development of science if the people on the one side who see more the turbulent whirl of material in all its colorfulness and vagueness, and those on the other side who love nice structural schemata would not polemized against each other but rather realize that the work of both is necessary for science."²⁴ Neurath (1941/1983, pp. 221-222; 1944, pp. 9-14) did not publish much of his polemics against semantics, and even those papers that appeared did not become influential: nonetheless Neurath was quite good in building and maintaining relationships, organizing events and institutions, so Carnap might have feared that Neurath's attacks against semantics could have influenced the movement behind the curtains.

Let's get back to the main line of argument of this paper. Besides the Tarski-style, referentially oriented version of semantics, there was another line of thought in Poland (Woleński, 2009): namely Leon Chwistek's approach, which was called in 1924 "Semeiotics" and, later, "Semantik" in 1932/1933 when his paper appeared in *Erkenntnis* and when Carnap mentioned it in *Syntax*. Chwistek's semantics was, however, quite different from Tarski's and from the Twardowski-school: it is best compared to Hilbert's theory of proof, to his formal method, since it disregards the relation of a sign to its reference; thus, as Carnap noted in LSL (§68), "Chwistek's system of so-called *semantics* is, on the whole, dedicated to the same task as our syntax" (emphasis in the original).

Again, what conclusion shall we draw now? Given the various terms and ranges of the scientific inquiry about language, meaning, and their relation to the extra-linguistic sphere, one shall distinguish two very general trends (painted with broad strokes): one type of semantics concerned the word-world relation, while the other treated the meaningful component, the information, and content-elements, which were conveyed through the use of signs. Notably, however, these trends were not always recognized or treated separately, or even if they were, the assessment of their value varied

²³ Carnap to Neurath, January 29, 1943 (RC 115-07-62).

²⁴ Carnap to Neurath, February 4, 1944 (RC 102-55-04).

quite frequently, not helping the general causes and aims of Carnap and other defenders of the new semantical project(s).

6. Carnap's usage of the terms

In order to draw conclusions about Carnap, we have to note whether there is any congruence in his and others' concept usage. As a start, let's take Carnap's official semantical writings. Both in *Foundations of Logic and Mathematics* (1939) and in *Introduction to Semantics* (1942), Carnap starts from the idea of semiotics as a *tripartite* inquiry of signs. The three levels are well known: pragmatics, semantics, and syntax. Pragmatics deals with signs and their users, in semantics we abstract from users and treat only the sign-world relation, and finally in syntax the relations of signs to signs become the primary target (Carnap, 1942, p. 9).

This approach, actually, corresponds nicely to Charles Morris's (1938) division of semiotic. But then we face some peculiarities! "With no real room for meanings under any label in [Morris's] system," says Burgess (ms.), "there is naturally no room in his *sogenant* semantics for anything like semantics as it had been conceived before World War I." Even though Morris became one of the main propagators of semantics in the United States and in (some parts of) Europe, his semantics was *not* the one which was already established outside of Poland. The same seems to be true for the mature Carnap as well, after he adopted Morris's program.

More importantly, however, Carnap's definition of semantics is very similar to that of Tarski. In his "The Concept of Truth in Formalized Languages," Polish edition 1933, German 1935, Tarski defined semantics as follows: "a characteristic feature of the semantical concepts is that they give expression to certain relations between the expressions of language and the objects about which these expressions speak" (Tarski, 1935/1956, p. 252). Interestingly though, while Carnap and Morris always started their discussion with the definition of 'semantics', Tarski defined it only at a very late point in the last part of his paper, thus somehow indicating that 'semantics' need not be defined (given presumably its well-known character).

Viewing matters from this perspective, it seems to be obvious that Carnap was under the philosophical and technical aegis of both Morris's and Tarski's semantics. It should be kept in mind, however, that there were important distinctions between Morris and Tarski. On November 22, 1936, Carnap visited Morris and talked with him about the *International Encyclopedia of Unified Science* and Tarski's semantics. He noted the following in his diaries: "Morris seems to be infected by Neurath's qualms, takes the notions, even my clarifications to be metaphysical, he

wants another — 'semantics' (?)."²⁵ Even though Morris fought for semantics, he was committed to behaviorism as well: just like Neurath. This relation is made explicit in Morris's controversial *Sign, Language, and Behavior*, where he said that "a science of signs can be most profitably developed on a biological basis and specifically within the framework of the science of behavior (a field which, following a suggestion of Otto Neurath, may be called *behavioristics*)" (Morris, 1946, p. 1). Note, however, that this was a quite typical move among American pragmatists; Nagel said in his Harvard lecture that

I can think of no better way to still these suspicions [about the metaphysical underpinning of semantics] than by placing the study of semantics into a behavioral context, and by instituting an analysis of such key semantic terms as 'designation' and 'truth' as used in *specific* contexts, in order to reveal the modes of action they signify. (Nagel, 1939/1954, p. 96. Emphasis in the original.)

So, semantics had the same agenda for Nagel and Morris as for Tarski and Carnap, but with different accounting: semantic notions shall be treated not as logical, but as behavioral concepts, thus freeing the project from actual and/or possible metaphysical components. Seemingly, there was as much agreement as disagreement among those who pursued semantics under the banner of logical empiricism.

It is not at all evident, after all, whether it was Morris or Tarski who had the most significant influence on Carnap after the mid-1930s. Given that Morris (1936) did not use 'semantics' or 'syntax' in his 1935 lecture at the International Congress for the Unity of Science in Paris,²⁶ but used instead the concept of 'existential dimension' and 'formal dimension of meaning', and given that Tarski presented his semantical ideas there in detail (invited by Carnap to do so), he might have influenced Morris somehow. At least Morris seems to have taken 'semantics' from Tarski and 'syntax' maybe from Carnap. In his unpublished lecture, entitled "Über den semantischen Wahrheitsbegriff" prepared for a discussion with Neurath in Paris, 1937, Carnap did not consider the tripartite approach of semiotic (so characteristic of Morris), or the peculiar logical ideas of Tarski. He was busy with the distinction of *truth* and *confirmation* (RC 080-32-01). Though the character of Carnap and Neurath's debate could explain this, it still documents Carnap's interests and concept-usage before his official semantical books.

Note also that although Tarski was a regular figure in Carnap's discussions (he even visited Carnap in Prague), the latter's famous "Truth

²⁵ Carnap's diary entry, November 22, 1936 (RC 025-82-01).

²⁶ The various events at and after the 1935 Paris Congress are described and analyzed from the viewpoint of the history of philosophy by Jan Woleński (forthcoming).

and Confirmation” paper (1949) — presented at the 1935 Paris Congress — mentions Tarski only once and doesn’t consider semantics *per se*, just the difference between truth and confirmation. What may cause some historical and interpretational trouble is that the 1949 English translation (prepared by “H.F.” which stands presumably for “Herbert Feigl”) is actually a mixture of the original “Wahrheit und Bewährung” paper and the 1946 “Remarks on Induction and Truth” paper (written for the discussion with Felix Kaufmann at a symposium) which deals in more detail with Tarski and semantics in the fashion and emphasis of the 1940s (Kaufmann presented Neurath-style arguments against Carnap’s semantics).

But anyway, what is certain is that Carnap thought that semantics should deal with signs, especially linguistic artifacts and their denotation or referents in the world. Later, in his posthumously published “Notes on Semantics,” Carnap (1972, p. 8) called his semantics “*pure, designative semantics*,” dealing with the “designative meaning component” which is “relevant for questions of truth.”

To move along, we shall ask the question whether this is the semantics that is to be found in Carnap’s LSL? Let’s see first what Carnap did not talk about. Obviously, he was not dealing with the referents, denotation, or, to be more precise, the *Bedeutung* of expressions — syntax shall abstract from these. Since ‘abstracting from the referents’ meant being formal, ‘formal semantics’ was just an oxymoron.

Interestingly, Carnap did not talk about the meaning, that is, the ‘Sinn’ of expressions. Or to be more precise, he did not consider meanings to be *separate, abstract entities in their own right*, to be grasped and expressed; and he did not take meaning to be such a function or mode of presentation that determines the *Bedeutung* — from this perspective, Frege was not an ideal in *Syntax*.

What Carnap did talk about is *content*, or in German, ‘Gehalt’/‘Inhalt’. He distinguished two components of expressions, namely Sinn/Gehalt/Inhalt and *Bedeutung*. Regarding their English translations, some confusion might arise. ‘Gehalt’ and ‘Inhalt’ is always translated as ‘content’; ‘Bedeutung’, however, is often rendered as ‘meaning’, but in many cases, ‘meaning’ is the word for ‘Sinn’, too. In other cases, ‘Sinn’ is just ‘sense’. As a result, often when Carnap speaks about how syntax abstracts from ‘meaning’, he means abstracting from ‘Bedeutung’, that is, the extra-linguistic element, not from ‘Sinn’, the content or information conveyed by expressions.

Logical syntax does not consider *Bedeutung*, but what is the case with ‘Sinn’, ‘Inhalt’ or ‘Gehalt’? Carnap asks whether it is possible to approach and define this component from the formal, or syntactical point of view, without taking into account the *Bedeutung*. His account of the content of a proposition is quite surprising (from the point of view of ORV, and even

URV) and interesting: the content of a proposition is what we can learn from it, what it conveys to us, that is, and here is the point, 'what we can deduce from S', and 'from what S follows.' 'Folgerung', 'entailment', 'consequence' are formal notions in the sense that they do not rely on Bedeutung. As Carnap (1934b, p. 11) says in "On the Character of Philosophic Problems" (written for the first issue of *Philosophy of Science*), "the question, whether a certain proposition is an inference [Folgerung] of certain other propositions or not, is therefore completely analogous to the question whether a certain position in chess can be played from another or not." Since content or Sinn is definable by Folgerung or consequence, it is also a *formal notion*. Actually, this is the main tenet of so-called inferentialism, as Jaroslav Peregrin (2014 and forthcoming) stated recently: the meaning of (logical and non-logical) expressions is exhausted by their inferential roles.

By the content of a sentence *S* Carnap meant the class of non-valid consequences of *S*, or the non-analytic consequences of *S*. Part 4 of Semantik's original manuscript contained a section entitled 'Theory of the Content of Formulas' with such subsections as 'The General Theory of Contents', 'Logical Content', 'Physicalistic Content', 'Empiristic Content,' and so on, but more importantly, its last section was: "Sinn = theoretischer Gehalt" (RC 110-04-07/p.3. IV/B/1-2). In fact, when he refers to his new "Attempt" manuscript in the "Overcoming Metaphysics" paper, Carnap (1932/1959, p. 62) says that "it is planned to give elsewhere a detailed exposition of metalogic as the theory of syntax and meaning [Sinnes], i.e. relations of deducibility [Ableitungsbeziehung]." This happened, after all, in the *Syntax* book, though under different names.

7. Concluding remarks

The aim of this paper was to show that defenders of ORV and URV are both right *in a sense*: Carnap indeed had a radical break in the mid-1930s, after the publication of his *Logical Syntax of Language*. What surfaces in his later semantical writings — the semantical concept of truth, based on denotation, reference, satisfaction etc. — is missing from LSL. But it must be admitted as well that Carnap already had some semantical insights in the context of the syntax project. These were partly connected to his later works (as it was emphasized by URV), but mainly they constituted semantics in a different sense: namely, as semantics dealing with content, information and with the 'Sinn' of expressions. Referentially conceived semantics as

a research project was pursued only later, and it is a further question whether Tarski or Morris represented the major line for Carnap.²⁷

Semantics for Carnap, in the early 1930s, was very similar to what later became called inferentialism, so, finally, I would like to point out some further historical and general ideas that may help us assessing Carnap's inferentialism, and provide perhaps further discussion.

Peregrin claimed (2014, pp. 27-28) that while Carnap accepted the usual empiricist dogma that the choice of conceptual frameworks or schemes is wholly conventional and arbitrary, and "the world can have its say only *after* we accomplish the choice," Sellars and others "believed that the world influences our putting together the framework" (emphasis in the original). But that is something entirely accepted by Carnap! In his "Theoretical Questions and Practical Decisions," published in the same year as *Syntax*, Carnap (1934c, p. 258) claimed that the world always has its say: while the development of linguistic frameworks, logical systems, and conceptual tools is a matter of convention, we are always restricted by already established empirical findings and by our wishes, aims, and goals pursued in a natural setting. In 1970, Carnap wrote the following to Gerhard Kroiss; it is worth quoting the letter at some length:

The question how and where we should begin with the construction of our language and our system of concepts is indeed an important problem and a rather puzzling one. Wherever we might begin, we feel the need for a clarification of our initial concepts, and this a need to use other concepts which precede the "initial" concepts. This is the old puzzle of philosophy: how and where should we make the absolute beginning?

I believe that the question in this form is unanswerable. It seems to me that the best description of the problem situation and of the required procedure was given by Neurath by his comparison with the task of repairing a ship. Our language, the whole conceptual system with which we wish to work, is like a ship on the open ocean; there is no dry-dock in which we could make a construction from an absolute beginning. We have only the ship itself, and we cannot proceed otherwise than to reconstruct parts while we are supported by the other parts. (Carnap to Gerhard Kroiss, February 5, 1970. RC 027-31-04. Typos corrected; pencil underlining by Carnap.)

Carnap thought (together with Neurath) that the world always has its say; we cannot free ourselves entirely to find an abstract realm with free-floating conventions. Carnap embraced — as André Carus (2007) called it — a "dialectical picture": setting up languages is a task for linguistic engineering and thus a practical task. How we set up our languages is based on how we set up our *rules* of formation and transformation. We may be

²⁷ Though Carnap always referred to Tarski, the tripartite conception of pragmatics, semantics, and syntax, and the conceptual settings of his works connect Carnap mainly to Morris. Furthermore, there are some interesting differences between Tarski and Carnap on which see Wagner 2017.

motivated by theoretical virtues and facts of the matter, but these facts shall be captured already in a certain framework. So theories (i.e., knowledge and facts) influence our practices (building up languages), but our practices (e.g. our decisions) influence the theoretical level (since you can know only those things that can be expressed in a language).

Peregrin (2014, p. 28) also says, that while for logical empiricists natural laws “must be simply empirical generalizations” and become like free-floating conventions, “Sellars noticed that, as a matter of fact, natural laws *do* take part in conferring roles on the terms they involve and hence that they do constitute concepts” (emphasis in the original). Though Carnap (1937, §82) accepts that we deal with conventions here, he says that they are not arbitrary. “The choice of them is influenced, in the first place, by certain practical methodological considerations [...],” but later “all must be tested by experience.” Of course, testing is not a unidirectional process, since after testing “P-rules [that is, natural laws] can be altered [...], or the protocol-sentence can be taken as being non-valid; or again the L-rules which have been used in the deduction can also be changed.” After all, for Carnap, the fact that “the hypotheses, in spite of their subordination to empirical control by means of the protocol-sentences, nevertheless contain a conventional element is due to the fact that the system of hypotheses is never univocally determined by empirical material, however rich it may be.”

Thus it is not entirely correct to say, as Peregrin does, that Carnap did not consider empirical languages and expression in his Syntax-project. It is true, that he focused on logic and mathematics, but his P-rules, functioning as transformation rules, were just *rules* that determined the meaning and content of physical expressions via *inferential patterns*. Rules, rules, and rules: in the semantical writings, Carnap claimed that “rules determine the *meaning* or *sense* of the sentences” (1942, §7; emphases in the original). Whether to accept a rule or to construct a new one is a question of *expedience*, as was everything for Carnap, including semantics and syntax (in whichever sense).

References

- Awodey, S. (2007). Carnap's Quest for Analyticity: the *Studies in Semantics*. In Friedman, M. & Creath, R. (eds.), *The Cambridge Companion to Carnap* (pp. 226-247). Cambridge: Cambridge University Press.
- Ayer, A. J. (1982). The Vienna Circle. In Gadol, E. T. (Ed.), *Rationality and Science. A Memorial Volume for Moritz Schlick in Celebration of the Centennial of His Birth* (pp. 36-54). Wien/New York: Springer.

- Black, M. (1945): Review of *Introduction to Semantics and Formalization of Logic*. *Mind*, 54 (214), 171-176.
- Burgess, J. (2008). Tarski's Tort. In *Mathematics, Models, and Modality. Selected Philosophical Essays*. (149-168). Cambridge: Cambridge University Press.
- Burgess, J. (ms.). Episodes in the History of 'Semantics.' Unpublished conference talk for the Princeton Conference: 'Semantics: Mathematics or Psychology?' September 15, 2013.
- Carnap, R. (1932). Erwiderung auf die vorstehenden Aufsätze von E. Zilsel und K. Duncker. *Erkenntnis*, 3, 177-188.
- Carnap, R. (1932/1959). The Elimination of Metaphysics Through Logical Analysis of Language. In Ayer, A. J. (Ed.), *Logical Positivism* (pp. 60-81). New York: The Free Press.
- Carnap, R. (1934a). *The Unity of Science*. (M. Black, Trans.) London: Kegan Paul. (Original work published in 1932.)
- Carnap, R. (1934b). On the Character of Philosophic Problems. *Philosophy of Science*, 1, 5-19.
- Carnap, R. (1934c). Theoretische Fragen und praktische Entscheidungen. *Natur und Geist*, 2 (9), 257-260.
- Carnap, R. (1937): *Logical Syntax of Language*. (A. Smeaton, Trans.) London: Routledge & Kegan Paul Ltd. (Original work published in 1934.)
- Carnap, R. (1949). Truth and Confirmation. In Feigl, H. & Sellars W. (Eds.), *Readings in Philosophical Analysis* (pp. 119-127). New York: Appleton-Century-Crofts, Inc.
- Carnap, R. (1950/1956). Empiricism, Semantics, and Ontology. Reprinted in *Meaning and Necessity* (pp. 205-221). Chicago: University of Chicago Press.
- Carnap, R. (1972). Notes on Semantics. *Philosophia*, 2, 3-54.
- Carus, A. (1999). Carnap, Syntax, and Truth. In Peregrin, J. (Ed.) *Truth and its Nature (If any)* (pp. 15-35). Dordrecht: Springer.
- Carus, A. (2007). *Carnap in Twentieth-Century Thought: Explication as Enlightenment*. Cambridge: Cambridge University Press.
- Coffa, A. (1976). Carnap's Sprachanschauung Circa 1932. *PSA*, 2, 205-241.
- Coffa, A. (1991). *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Cambridge: Cambridge University Press.
- Creath, R. (1990a). The Unimportance of Semantics. *PSA*, 2, 405-416.
- Creath, R. (1990b). (Ed.). *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work*. Berkeley: University of California Press.
- Creath, R. (1996). Languages Without Logic. In Giere, R. N. & Richardson, A. W. (eds.) *Origins of Logical Empiricism* (pp. 251-265). Minneapolis: University of Minnesota Press.

- Creath, R. (1999). Carnap's Move to Semantics: Gains and Losses. In Woleński, J. & Köhler, E. (eds.) *Alfred Tarski and the Vienna Circle – Austro-Polish Connections in Logical Empiricism* (pp. 65-76). Dordrecht: Kluwer Academic Publishers.
- Gruber, M. (2016). *Alfred Tarski and the "Concept of Truth in Formalized Languages." A Running Commentary with Consideration of the Polish Original and the German Translation*. Dordrecht: Springer.
- J. M. R. (1942). Review of Carnap's *Introduction to Semantics*. *Philosophy of Science*, 9 (3), 281-282.
- Kokoszyńska, M. (1936). Syntax, Semantik und Wissenschaftslogik. *Actes du Congrès International de Philosophie Scientifique, fasc. 3* (pp. 9-14). Paris: Hermann & Édituers.
- Kokoszyńska, M. (1939). Review of Carnap's *Foundations of Logic and Mathematics*. *The Journal of Symbolic Logic*, 4 (3), 117-118.
- Morris, Ch. (1936). The Concept of Meaning in Pragmatism and Logical Positivism. *Actes du Congrès International de Philosophie Scientifique, fasc 8* (pp. 130-138). Paris: Hermann & Édituers.
- Morris, Ch. (1946). *Sign, Language, and Behavior*. New York: Prentice Hall.
- Nagel, E. (1935). Review of *Logische Syntax der Sprache*. *Journal of Philosophy*, 32, 49-52.
- Nagel, E. (1939/1954). Charles S. Peirce, Pioneer of Modern Empiricism. In *Sovereign Reason and Other Studies in the Philosophy of Science* (pp. 89-100). Glencoe, Illinois: The Free Press.
- Nagel, E. (1942). Review of *Introduction to Semantics*. *The Journal of Philosophy*, 39 (17), 468-473.
- Neurath, O. (1941/1983). Universal Jargon and Terminology. In Cohen, R. S. & Neurath, M. (Eds.) *Otto Neurath: Philosophical Papers 1913-1946* (pp. 213-229). Dordrecht: D. Reidel.
- Neurath, O. (1944). *Foundations of Social Sciences*. Chicago: University of Chicago Press.
- Partridge, E. (1945). Review: *Introduction to Semantics* by Rudolf Carnap. *Journal of the Royal Society of Arts*, 93, 455-456.
- Peregrin, J. (2014). *Inferentialism. Why Rules Matter*. Basingstoke: Palgrave Macmillan.
- Peregrin, J. (forthcoming). Carnap's Inferentialism. In Schuster, R. (ed.) *Vienna Circle in Czechoslovakia*. Dordrecht: Springer.
- Read, A. W. (1948). An Account of the Word 'Semantics.' *WORD*, 4 (2), 78-97.
- Reisch, G. (2005). *How the Cold War Transformed Philosophy of Science: to the Icy Slopes of Logic*. Cambridge/New York: Cambridge University Press.

- Ricketts, Th. (1996). Carnap: From Logical Syntax to Semantics. In Giere, R. N. & Richardson, A. W. (eds.) *Origins of Logical Empiricism* (pp. 213-250). University of Minnesota Press.
- Tarski, A. (1935/1956). The Concept of Truth in Formalized Languages. In Tarski, A. *Logic, Semantics, Metamathematics. Papers from 1923 to 1938* (pp. 152-278). Oxford: Clarendon Press.
- Tarski, A. (1936/1956). On the Concept of Logical Consequence. In *Logic, Semantics, Metamathematics. Papers from 1923 to 1938* (pp. 409-420). Oxford: Clarendon Press.
- Wagner, P. (2017). Carnapian and Tarskian Semantics. *Synthese*, 194 (1), 97-119.
- Woleński, J. (1999). Semantic Revolution: Rudolf Carnap, Kurt Gödel, Alfred Tarski. In Woleński, J. & Köhler, E. (eds.) *Alfred Tarski and the Vienna Circle* (pp. 1-15). Dordrecht: Kluwer Academic Publishers.
- Woleński, J. (2009). The Rise and Development of Logical Semantics in Poland. In Lapointe, S. & Woleński, J. & Marion, M. & Miskiewicz, W. (eds.) *The Golden Age of Polish Philosophy: Kazimierz Twardowski's Philosophical Legacy* (pp. 43-59). Springer.
- Woleński, J. (2017). The Story of the German Translation of Tarski's Book of Truth. In Bar-Am, N. & Gattei, S. (eds.) *Encouraging Openness. Essays for Joseph Agassi on the Occasion of His 90th Birthday* (pp. 81-90). Dordrecht: Springer.
- Woleński, J. (forthcoming). On Controversy on Semantics at the Paris Congress 1935.

